

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

cleic - nucleic search, using sw model

December 9, 2003, 01:21:55 : Search time 2842 Seconds

(without alignments)
9817.159 Million cell updates/sec

US-10-035-914-1

1 aaatccgcctctcgtctc.....gaacagctcgtgactgag 682

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

2888711 seqs, 2045481386 residues

number of hits satisfying chosen parameters: 5777422

DB seq length: 0
DB seq length: 2000000000

Maximum Match 0%

Listing first 45 summaries

use :
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rpd:*
36: em_hcg_vit:*
37: em_hcg_vit:*
38: em_hcg_hum:*
39: em_hcg_mus:*
40: em_hcg_mus:*
41: em_hcg_other:*

Result
No. Score Match Length DB ID

Result No.	Score	Match	Length	DB	ID	Description
1	682	100.0	1234	10	MMMD3MRN	X59047 M. musculus
2	682	100.0	1510	10	BC011433	BC011433 Mus muscu
3	614.8	90.1	1303	6	AX700189	AX700189 Sequence
4	614.8	90.1	1303	10	RNU19894	U19894 Rattus norv
5	535.4	78.5	711	9	AF116599	AF116599 Chloroceb
6	534.8	78.4	1333	9	BC002978	BC002978 Homo sapi
7	534.8	78.4	1496	6	AX337094	AX337094 Sequence
8	534.8	78.4	1496	6	AX658153	AX658153 Sequence
9	534.8	78.4	1496	9	HUMTAPAL	M13680 Human 26-kD
10	530.6	77.8	711	9	SOE250197	AJ250197 Saguinus
11	527.4	77.3	711	9	AF116600	AF116600 Pan trogl
12	484.2	71.0	633	9	AF274885	AF274885 Saguinus
13	290.4	42.6	1329	5	BC041217	BC041217 Xenopus 1
14	277.4	40.7	1300	5	AF295377	AF295377 Dario rer
15	270	39.6	436	6	AR240547	AR240547 Sequence
16	219.2	32.1	429	6	AX198983	AX198983 Sequence
17	217.8	31.9	316	6	AX209509	AX209509 Sequence
18	213.4	31.3	311	6	AR272638	AR272638 Sequence
19	213.4	31.3	311	6	AR276219	AR276219 Sequence
20	213.4	31.3	311	6	AX062740	AX062740 Sequence
21	213.4	31.3	311	6	AX367657	AX367657 Sequence
22	195.2	28.6	357	6	AX004409	AX004409 Sequence
23	195.2	28.6	357	6	BD077115	BD077115 Hepatitis
24	190.8	28.0	348	6	AX004411	AX004411 Sequence
25	190.8	28.0	348	6	BD077116	BD077116 Hepatitis
26	184	27.0	1513	5	AY152677	AY152677 Petromyzo
27	165	24.2	1240	4	BOVCD9ANT	M61720 Bos taurus
28	155.4	22.8	1227	4	AY072785	AY072785 Sus scrof
29	153.8	22.6	1024	4	CATPCD9	D30786 Felis catus
30	153.8	22.6	1150	4	CATCDAN	L35275 Felis catus
31	153	22.4	1054	4	AF525029	AF525029 Sus scrof
32	152.6	22.4	687	6	EO5732	EO5732 cDNA encodi
33	151	22.1	687	6	AR016440	AR016440 Sequence
34	151	22.1	687	6	113743	113743 Sequence 2
35	151	22.1	687	6	HUMP24CD9A	L34068 Human antlg
36	151	22.1	1120	6	AR016441	AR016441 Sequence
37	151	22.1	1120	6	BD176845	BD176845 A method
38	151	22.1	1120	9	AGMDRAP27	D10726 African gre
39	151	22.1	1120	6	HSMPR1	X60111 H. sapiens m
40	151	22.1	1192	6	AX409704	AX409704 Sequence
41	151	22.1	1192	6	AX658227	AX658227 Sequence
42	151	22.1	1192	9	HUMANTCD9	M38690 Human CD9 a
43	151	22.1	66877	2	AC116062	AC116062 Rattus no
44	149.4	21.9	1120	6	113744	113744 Sequence 3
45	149.4	21.9	1166	9	BC011988	BC011988 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS MMMD3MRN 1234 bp mRNA linear ROD 27-OCT-1992
DEFINITION M. musculus MD3 mRNA.
ACCESSION X59047
VERSION X59047.1 GI:53037
KEYWORDS cell surface antigen.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1234)
AUTHORS Duff, K. and Parsons, J.
TITLE Mouse Homologue of human TAPA-1 mRNA
JOURNAL Unpublished

gene /lab host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1510
/gene="Cd81"
/note="synonym: CD81"
/db_xref="LocusID:12520"
/db_xref="MGI:1096398"
207. 917
/codon_start=1
/product="CD 81 antigen"
/protein_id="AAH1433.1"
/db_xref="GI:15030313"
/db_xref="LocusID:12520"
/translation="MGVEGCTKCIKYLFFVNFVFLAGVILGLVALMLRHPDPTSL
LYLEAGNKPAPNPFYVGIYILAVGAVMFWPLGICGVAIOESOLLGPTFCTLVLF
ACEVAAGIWPVNDQIADKVKOPYDALQAVDDANNAKAVYKTHETLNCSSN
ALTLVTIILNSICPSGAILTFLLQDDCHQKIDELPSKLYLIGIALVAVIMIF
EMLSWLCGGINSSVY"

CDS
COUNT 340 a 400 c 360 g 410 t
N

100.0%; Score 682; DB 10; Length 1510;
Local Similarity 100.0%; Pred. No. 9.2e-182;
ches 682; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AAATACCTGCTCTGCTCTCAATTTGCTCTGCTGCTGAGGCGTATCTAGT 60
237 AAATACCTGCTCTGCTCTCAATTTGCTCTGCTGCTGAGGCGTATCTAGT 296
61 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
297 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 356
121 AACAAACCGGACCCCAACACCTCTACGTGCGATCTACATTTCTCATGCTGGAGCT 180
357 AACAAACCGGACCCCAACACCTCTACGTGCGATCTACATTTCTCATGCTGGAGCT 416
181 GTAGATGTTTGTAGGCTCTGAGGCTGCTATGAGGCGATCCAGAGTCCAGTCTG 240
417 GTATGATGTTTGTAGGCTCTGAGGCTGCTATGAGGCGATCCAGAGTCCAGTCTG 476
241 CTGGGACGTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
477 CTGGGACGTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 536
301 TGGGGCTTGTAAACAAGACGATGCGCAAGATGTGAAGCAATTCTATGACAGGCC 360
537 TGGGGCTTGTAAACAAGACGATGCGCAAGATGTGAAGCAATTCTATGACAGGCC 596
361 CTTCAGCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
597 CTTCAGCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 656
421 CATGAGACGCTCAACTGTTGTGCTCAACGCACTGACCACTGACCTACCACTACG 480
657 CATGAGACGCTCAACTGTTGTGCTCAACGCACTGACCACTGACCTACCACTACG 716
481 AGGAACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
717 AGGAACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
541 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
777 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 836
601 GTGGTAGCTGTCAATATGATCTTTGATGATGATGATGATGATGATGATGATGATGAT 660
837 GTGGTAGCTGTCAATATGATCTTTGATGATGATGATGATGATGATGATGATGATGAT 896
661 CGGAACAGCTCGGTATGAG 682
897 CGGAACAGCTCGGTATGAG 918

RESULT 3
AX700189 1303 bp DNA linear PAT 03-Apr-2003
Sequence 75 from Patent EP1264298.
AX700189
AX700189.1 GI:29536100
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 Brooksbank, R.A., Dixon, A.K., Lee, K. and Pinnock, R.D.
Identification and use of molecules implicated in pain
Patent: EP 1264298-A 75 19-PEB-2003;
WARNER-LAMBERT COMPANY (US)
FEATURES
Location/Qualifiers
1. 1303
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/note="Cd81 antigen"

BASE COUNT 257 a 375 c 347 g 324 t
ORIGIN

Query Match 90.1%; Score 614.8; DB 6; Length 1303;
Best Local Similarity 93.8%; Pred. No. 9.6e-163;
Matches 640; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
1 AAATACCTGCTCTGCTCTCAATTTGCTCTGCTGCTGAGGCGTATCTAGT 60
253 AAATACCTGCTCTGCTCTCAATTTGCTCTGCTGCTGAGGCGTATCTAGT 312
61 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
313 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
121 AACAAACCGGACCCCAACACCTCTACGTGCGATCTACATTTCTCATGCTGGAGCT 180
373 GACAAACGACGACCTGACCTCTTATGAGGCACTTACATTTCTCATGCTGGAGCT 432
181 GTAGATGTTTGTAGGCTCTGAGGCTGCTATGAGGCGATCCAGAGTCCAGTCTG 240
433 GTAGATGTTTGTAGGCTCTGAGGCTGCTATGAGGCGATCCAGAGTCCAGTCTG 492
241 CTGGGACGTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
493 CTGGGACGTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
301 TGGGGCTTGTAAACAAGACGATGCGCAAGATGTGAAGCAATTCTATGACAGGCC 360
553 TGGGGCTTGTAAACAAGACGATGCGCAAGATGTGAAGCAATTCTATGACAGGCC 612
493 CTTCAGCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
613 CTTCAGCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 672
421 CATGAGACGCTCAACTGTTGTGCTCAACGCACTGACCACTGACCTACCACTACG 480
673 CATGAGACGCTCAACTGTTGTGCTCAACGCACTGACCACTGACCTACCACTACG 732
481 AGGAACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
733 AGGAACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
541 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
793 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
601 GTGGTAGCTGTCAATATGATCTTTGATGATGATGATGATGATGATGATGATGATGAT 660

853 GTGTAGCTGTCATTATGATCTTCAGAGATGATTCGACGATGCTGCTGTGGCATC 912
661 CGGAACAGCTCCGTGACTAG 682
913 CGGAACAGCTCCGTGACTAG 934
1 4
94
TION RNU19894 1303 bp mRNA linear ROD 25-MAR-1997
Rattus norvegicus target of the antiproliferative antibody mRNA,
complete cds.
ION U19894
DS U19894.1 GI:1142641
Rattus norvegicus (Norway rat)
Rattus norvegicus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
NCE 1 (bases 1 to 1303)
ORS Geisert, E. B. Jr., Murphy, T. P., Irwin, M. H. and Larjava, H.
E A novel cell adhesion molecule, G-CAM, found on cultured rat glia
Neurosci. Lett. 133 (2), 262-266 (1991)
INE 92278609
MED 1816505
NCE 2 (bases 1 to 1303)
ORS Irwin, M. H. and Geisert, E. B. Jr.
E The upregulation of a glial cell surface antigen at the astrocytic
scar in the rat
Neurosci. Lett. 154 (1-2), 57-60 (1993)
NAL Neurosci. Lett. 154 (1-2), 57-60 (1993)
INE 91368802
MED 8361649
NCE 3 (bases 1 to 1303)
ORS Geisert, E. B. Jr., Yang, L. and Irwin, M. H.
E Astrocyte growth, reactivity, and the target of the
antiproliferative antibody, TAPA
J. Neurosci. 16 (17), 5478-5487 (1996)
NAL 96346153
INE 8757260
MED 4 (bases 1 to 1303)
NCE Geisert, E. B.
ORS Direct Submission
E Submitted (17-JAN-1995) Eldon B. Geisert Jr., Department of Anatomy
and Neurobiology, University of Tennessee, Memphis, 855 Monroe
Ave., Memphis, TN 38163, USA
Location/Qualifiers
BS 1. 1303
SOURCE /organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="1.1, sp1, sp2"
/sex="Male"
/cell_type="astrocyte (Type 1 and Type 2)"
/tissue_type="cerebral cortex"
/clone_id="Astrocyte cDNA lambda gt11 (Clontech No.
RL1042B)"
1. 222
223. 993
/note="TAPA"
/codon_start=1
/product="target of the antiproliferative antibody"
/protein_id="AAC5103.1"
/db_xref="GI:1142642"
/translat="MGVRCCKKIKYLLFVNPVPMLAGVILGVALLRDPOTTL
LYLHGDPASTPTVGIYIILAGVAMVFGFAGCTGAIQESQCLGFTCLVILF
ACEVAGIWMGVNDQAKDVKOPFDQALQAVMDNDANNAKAVKTHETLANCGSN
TLTTLTAVALNSLCPSSNSFTQLKEDCHKIDELFSGLYLIGIAIVAVAMIF
EMILSMVLCGIRNSSVY"
934. 1303
UTR 257 a 375 c 347 g 324 t
OUNT

Query Match 90.1%, Score 614.8; DB 10; Length 1303;
Best Local Similarity 93.8%; Pred. No. 9,66-163;
Matches 640; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
OY 1 AATACCTGCTCTTCTGCTTCAATTTGCTTCTTCTGCTGCTGAGCGGTATCTAGT 60
DB 253 AATACCTGCTCTTCTGCTTCAATTTGCTTCTTCTGCTGCTGAGCGGTATCTAGT 312
OY 61 GTAGCTGTGTGGTGGCGCATGATCCAGACGACGCTGTGTACTGTGAATCTGGA 120
DB 313 GTAGCTGTGTGGTGGCGCATGATCCAGACGACGCTGTGTACTGTGAATCTGGA 372
OY 121 AACAAACCGGACCCCAACACCTTCTAGCGGGCATCTACATTTCTCATTTGCTGGA 180
DB 373 GACAAACGACGACCTTCTAGCATCTTCTATGTGGGCACTTACATTTCTCATTTGGA 432
OY 181 GTGATGATGTTTGAAGCTTCTGAGGCTGTAGAGGCGCATCCAGAGTCCAGTGTCTG 240
DB 433 GTGATGATGTTTGAAGCTTCTGAGGCTGTAGAGGCGCATCCAGAGTCCAGTGTCTG 492
OY 241 CTGGGGACGTTCTTCACTGCTGTGTATCTGTTTCTGTGTAGGTGTGTGAGGATC 300
DB 493 CTGGGGACGTTCTTCACTGCTGTGTATCTGTTTCTGTGTAGGTGTGTGAGGATC 552
OY 301 TGGGGCTTCTTAAACAAAGACAGATCGCCAAAGATGAAAGAGTTCTATAGACAGGCG 360
DB 553 TGGGGCTTCTTAAACAAAGACAGATCGCCAAAGATGAAAGAGTTCTATAGACAGGCG 612
OY 361 CTTCAGCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
DB 613 CTTCAGCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 672
OY 421 CATGAGAGCTTCACTGTTGTGCTTCAAGGACGACGACGACGACGACGACGACGAC 480
DB 673 CATGAGAGCTTCACTGTTGTGCTTCAAGGACGACGACGACGACGACGACGACG 732
OY 481 AGGAACAGCTGTGCTTCTGAGGCGGCAACATCTCAACCCCTTCTAGTGTGAGATG 540
DB 733 AGGAACAGCTGTGCTTCTGAGGCGGCAACATCTCAACCCCTTCTAGTGTGAGATG 792
OY 541 CATCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 793 CATCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
OY 601 GTGTAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 853 GTGTAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 912
OY 661 CGGAACAGCTCCGTGACTAG 682
DB 913 CGGAACAGCTCCGTGACTAG 934
RESULT 5
AF116599 711 bp mRNA linear PRI 01-FEB-1999
LOCUS Chlorocebus aethiops CD81 mRNA, complete cds.
DEFINITION AF116599
ACCESSION AF116599.1 GI:4206170
KEYWORDS
SOURCE
ORGANISM
Cercopithecus aethiops (African green monkey)
Cercopithecus aethiops
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Cercopithecus.
REFERENCE
1 (bases 1 to 711)
AUTHORS Levy, S. and Kuo, C. C.
TITLE African green monkey CD81 cDNA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 711)
AUTHORS Levy, S. and Kuo, C. C.
TITLE Direct Submission

	BASE COUNT	257 a	504 c	413 g	322 t	/db_xref=taxon:9606"
	ORIGIN					
	Query Match	78.4%	Score 534.8;	DB 6;	Length 1496;	
	Best Local Similarity	86.5%;	Pred. No. 4.4e-140;			
	Matches 590;	Conservative	0;	Mismatches 92;	Indels 0;	Gaps 0;
Dy	1	AAATACCTGCTCCTTCGTCTTCAATTGGTTCTTGCGCTGGAGCGTGAATCCCTTAGT 60				
Dd	269	AAGTAAGCTGCCTCTTCGCTTCATATTTGCTCTTGCTGGCTGGAGCGTAATCTTGACT 328				
Dy	61	GTAGCTCTGTGTGTTGCGTCAATGATCCAAGAACAAGAGCTGTTACTTGAACCTGGA 120				
Dd	329	GTGGCCCTGTGGCTCCGCCATGACCAGCACCAAACCTCCTGTATCTGAGAGCTGGGA 388				
Dy	121	AACAAACCGGACCCCAACACCTTCTGAGTGGGCACTTACATTCTCATTTGCTGTGGAGCT 180				
Dd	389	GACAAAGCCGCGCCCAACACCTTCTAATGAGGCAATCAATCCCTCAATCGCTGTGGGCGCT 448				
Dy	181	GTGATGATGTTTTGAGGCTTCTGGGGGTGCTATGAGGGCCATCCAGAGAGTCCCAGTGTCTG 240				
Dd	449	GTATATATTTGCTGTGGCTTCTTGAGCTGTGACGGGCCATCCAGAAATCCCAATGTCTG 508				
Dy	241	CTGGGAGCGTTCTTCAACCTGCTGTGATCTGTTTGCCTGTGAAGTGTGGCTGACAGCATC 300				
Dd	509	CTGGGAGCGTTCTTCAACCTGCTGTGATCTGTTTGCCTGTGAAGTGTGGCTGACAGCATC 568				
Dy	301	TGGGGCTTTCGTAAACAAGACCAAGATGCGCAAGATGTGAAGCATTTCTATGACCAAGGCC 360				
Dd	569	TGGGGCTTTCGTAAACAAGACCAAGATGCGCAAGATGTGAAGCATTTCTATGACCAAGGCC 628				
Dy	361	CTTCAGGACAGCTGTGATGATGATGATGATGCCAAGAGCGGAGCTGTGGTGAAGACTTTC 420				
Dd	629	CTACAGGAGGCGCGTGGTGAATGATGATGATGCCAAGAGCGGAGCTGTGGTGAAGACTTTC 688				
Dy	421	CATGAGACGCTCAACTGTGTGTGGCTCCAGCACTGAACAACAATACTGACCAACTATCTG 480				
Dd	689	CAGAGACGCTGTACCTGT 748				
Dy	481	AGGACAGCGCTGT 540				
Dd	749	AAGAACATTTGT 808				
Dy	541	CATCAGAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600				
Dd	809	CACCAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 868				
Dy	601	GTGGTAGCTGTCAATTAATGATCTTTGAGATGATTTCTGAGCATGTGGTGTGTGTGTGTGTGT 660				
Dd	869	GTGGTAGCTGTCAATTAATGATCTTTGAGATGATTTCTGAGCATGTGGTGTGTGTGTGTGTGT 928				
Dy	661	CGAAGACGCTTCGCTGATCTGAG 682				
Dd	929	CGAAGACGCTTCGCTGATCTGAG 950				
RESULT 8						
LOCUS	AX658153	1496 bp	DNA	linear	PAT 22-MAR-2003	
DEFINITION	Sequence 69 from Patent WO03000928.					
ACCESSION	AX658153					
VERSION	AX658153.1	GI:29160735				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Bukaryoga; Metazoa; Chordata; Craniata; Vertebrata; Rutelesotomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
AUTHORS	Poulsen,H.S., Pedersen,N., Mortensen,S., Sorensen,S.B., Petersen,M.W. and Elner,H.I.					
TITLE	Methods for identification of cancer cell surface molecules and cancer specific promoters, and therapeutic uses thereof					

809 CACGAGAGATGATGATGATCTTCTTCGCGGAGAGCTGATCTCATGCGATTCGTCGATC 868
 601 GTGGATCTGTCAATATATATCTTTGAGATGATTCGACGATGTCGTCGTCGTCGATC 660
 869 GTGGTCGCTGTGATCATATCTTCGAGATGATCTCGACATGTCGTCGTCGTCGATC 928
 661 CGGAACAGCTCCGCTACTGAG 682
 929 CGGAACAGCTCCGCTACTGAG 950

10
 197
 TION SOR250197 711 bp mRNA linear PRI 28-JUN-2000
 ION Saguinus oedipus mRNA for CD81 protein (cd81 gene).
 N AJ250197.1 GI:7688356
 DS cd81 gene; CD81 protein.

NISM Saguinus oedipus (cotton-top tamarin)
 Saguinus oedipus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
 Saguinus.

NCE 1
 Meola, A., Sbardellati, A., Bruni, E.B., Cerretti, M., Pezanes, M.,
 Cecacci, A., Vitelli, A., Levy, S., Nicotia, A., Traboni, C.,
 McKeating, J. and Scarselli, E.
 Binding of hepatitis C virus E2 glycoprotein to CD81 does not
 correlate with species permissiveness to infection

8 J. Virol. 74 (13), 5933-5938 (2000)

NAL J. Virol. 74 (13), 5933-5938 (2000)
 MED 10846074
 NCE 2 (bases 1 to 711)
 JRS Traboni, C.
 S Direct Submission
 Submitted (08-OCT-1999) Traboni C., Biochemistry, IRBM P.
 Angeletti, via Pontina km. 30.600, 00040 Pomezia (Roma), ITALY
 SS Location/Qualifiers
 1. 711

/organism="Saguinus oedipus"
 /mol_type="mRNA"
 /db_xref="taxon:9490"
 /clone="CD81. ORF.S02"
 /tissue_type="liver"
 1. 711
 /gene="cd81"
 1. 711
 /gene="cd81"
 /codon_start=1
 /product="CD81 protein"
 /protein_id="CA89875.1"
 /db_xref="GI:7688357"

/translation="MGVEGCTKCIKYLFPVNFVFMAGVILIGVALMLRHDSQNTL
 LYELEGRKPAAPNTFYVGIYILIVAGVMMVPRGCGAIGESQCLGTFTEVLVLF
 ACERVAATIKGVKQDAKDYKQTDALQAAVDDANNAKAVKTFHEHLCSS
 TLSALWTLMLNNLCPSSGSSITISLRFEDCHKIDELFSGKLYLIGALVAIVIMIF
 EMLISWLCGIRNSYV"
 JUNT 132 a 213 c 205 g 161 t

/ Match 77.8%; Score 530.6; DB 9; Length 711;
 Local Similarity 86.2%; Pred. No. 6.2e-139;
 nes 587; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

1 AATATCTGCTCTGCTCTTCAATTTGCTTTGCGGCTGAGGCGTATCTTAGT 60
 31 AAGTACTGCTCTTCTGCTTCACTTCTTCTGCGCTGAGGCGTATCTTAGT 90
 61 GTAGCTGTGCTGCTGCTATGATCAACAACAAGCTCTGTACTGTGAATGGGA 120
 91 GTGGCCCTGTGGCTCCGCAATGACCCGCAACCAACCTCTGTATCTGAGCTGGGA 150

QY 121 AACAAACCGGACCCCAACACCTTTATAGTGGGATCTCAATTCATTCGTGGAGCT 180
 Db 151 GACAGGCTGCGGCCCAACCTTTCTAGTAGGATCTCAATTCATTCGTGGAGCT 210
 QY 181 GTATGATGTTTATAGGCTTCGAGGGGTGTATAGGGCCATCCAGAGTCCAGTGTG 240
 Db 211 GTATGATGTTTATAGGCTTCGAGGGGTGTATAGGGCCATCCAGAGTCCAGTGTG 270
 QY 241 CTGGGAGCTTTCTTCACTGCTGCTGTGATCTGTTGCTGTGAGGTGCTGGAGATC 300
 Db 271 CTGGGAGCTTTCTTCACTGCTGCTGTGATCTGTTGCTGTGAGGTGCTGGAGATC 330
 QY 301 TGGGCTTCTTAAACAAGACAGATGCGCAAGATGTGAAGAGTTCTATAGCAAGGCC 360
 Db 331 TGGGCTTCTTCAACAAGACAGATGCGCAAGATGTGAAGAGTTCTATAGCAAGGCC 390
 QY 361 CTTCAGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
 Db 391 CTGACAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 450
 QY 421 CATGAGAGCTCAACTGTTGTGCTTCAAGCACTGACCACTGACCACTGACCACTGAC 480
 Db 451 CAGAGAGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
 QY 481 AGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 511 AAGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 570
 QY 541 CACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 Db 571 CACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
 QY 601 GTGTAGCTGTCAATATATATCTTTGAGATGATTCGAGAGTGTGCTGTCGTCGATC 660
 Db 631 GTGTAGCTGTCAATATATATCTTTGAGATGATTCGAGAGTGTGCTGTCGTCGATC 690
 QY 661 CGGAACAGCTCCGCTACTGAG 681
 Db 691 CGGAACAGCTCCGCTACTGAG 711

RESULT 11.
 AF116600 711 bp mRNA linear PRI 01-FEB-1999
 LOCUS Pan troglodytes CD81 mRNA, complete cds.
 DEFINITION AF116600
 ACCESSION AF116600
 VERSION AF116600.1 GI:4206172
 KEYWORDS Pan troglodytes (chimpanzee)
 SOURCE Pan troglodytes
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.

REFERENCE 1 (bases 1 to 711)
 Levy, S. and Kuo, C.C.
 CD81 cDNA sequence derived from chimpanzee cells
 JOURNAL Unpublished
 TITLE CD81 cDNA sequence derived from chimpanzee cells
 AUTHORS 2 (bases 1 to 711)
 REFERENCES Levy, S. and Kuo, C.C.
 JOURNAL Direct Submission
 TITLE Submitted (24-DEC-1998) Medicine/Oncology, Stanford, Stanford, CA
 JOURNAL 94305, USA

FEATURES
 source Location/Qualifiers
 1. 711

/organism="Pan troglodytes"
 /mol_type="mRNA"
 /db_xref="taxon:9598"
 /cell_line="Janice"
 /note="BBV-transformed cells; amplified using Homo sapiens
 CD81-derived primers"
 1. 711
 /codon_start=1
 /product="CD81"
 /protein_id="AA011440.1"

CDS

```

/db xref="GI:4206173"
/translation="MGVEGCTCICIKLLEFVNFVPLAGVILGVALMLRHPDPTNL
LYLEIGDPRAPNTFYVGIYILAVGAVMNFVFLCYCAIOBSOCLTFPTCLILP
ACRVAAGWGPVNRHQAIDVQFIDOLQOAVDDNNANAVVTHETLNCSS
TITLITATVLRKNLCPSSGNNISNLFKEDCHOKIDLPFGKLYLIGIAIVAVIMIP
EMILSMVLCGGRNSVY"
primer_bind
primer_bind
COUNT 136 a 212 c 200 g 163 t
complement(693..-5711)

```

```

1 AAATACCTGCTCTTGTCTTCAATTTCTTCTGCTGCTGAGGCTGATCTAGT 60
31 AAGTACCTGCTCTTGTCTTCAATTTCTTCTGCTGCTGAGGCTGATCTAGT 90
61 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
91 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150
121 AACAAACCGGACCCCAACACCTTCTACGCTGAGGCTGCTGCTGCTGCTGCT 180
151 GACAAAGCCCGGACCCCAACACCTTCTAGGCTGCTGCTGCTGCTGCTGCTGCT 210
181 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
211 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 270
241 CTGCGGAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
271 CTGCGGAGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 330
301 TGCGGCTTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
331 TGCGGCTTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 390
361 CTGCGGAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
391 CTGCGGAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 450
421 CATGAGAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
451 CATGAGAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
481 AGGAAAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
511 AAGAAAGCTTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
541 CATGAGAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
571 CATGAGAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
601 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
631 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 690
661 CGGAAAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 690
691 CGGAAAGCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 711

```

```

1 12
985 AF274885 633 bp mRNA linear PRI 14-NOV-2000
ITON Saginus oedipus CD81 mRNA, partial cds.
310N AF274885
3N AF274885.1 GI:11141540
3DS
3 Saginus oedipus (cotton-top tamarin)
3N Saginus oedipus
3N Saginus oedipus

```

```

FEATURES
source
1.633
/organism="Saguinus oedipus"
/mol_type="mRNA"
/db_xref="taxon:9490"
<1..>633
/codon_start=1
/product="CD81"
/db_xref="GI:11141541"
/db_xref="L:11141541"
/translation="LFFVNFVPLAGVILGVALMLRHPDPTNLYLEIGDPRAPNT
FYVGIYILAVGAVMNFVFLCYCAIOBSOCLTFPTCLILPACRVAAGWGPVNR
HQAIDVQFIDOLQOAVDDNNANAVVTHETLNCSSSTLSLTMKLN
LCPSSGNNISNLFKEDCHOKIDLPFGKLYLIGIAIVAVIMIPFEMILSM"
BASE COUNT
116 a 195 c 178 g 144 t
ORIGIN
Query Match 71.0%; Score 484.2; DB 9; Length 633;
Best Local Similarity 85.3%; Pred. No. 8.4e-126;
Matches 540; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

```

```

10 CTCTGCTGCTTCAATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69
1 CTCTGCTGCTTCAATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
70 TGCTGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
61 TGCTGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
130 GCACCAACACCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 189
121 GCGCCCAACCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
190 TTGTGAGCTTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
181 TTGTGAGCTTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
250 TTGTGAGCTTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
241 TTGTGAGCTTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
310 GTAAACAAAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
301 GTAAACAAAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
370 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
361 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
430 CTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
421 CTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
490 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
481 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
550 ATCGATGAGCTTCTTCTGCGAAGCTGTAACCTGTAATGGAATTGCAAGCA 609

```


ABT07141 standard; cDNA; 436 BP.
ABT07141.

07-NOV-2002 (first entry)

Human ovarian cancer associated coding sequence SEQ ID NO: 303.

Human; ovarian cancer; cancer; gene; ss.

Homo sapiens.

US2002076715-A1.

20-JUN-2002.

06-JUN-2001; 2001US-0876889.

23-SEP-1998; 98US-0159320.
08-FEB-1999; 99US-0246429.
16-SEP-1999; 99US-0397787.

(BENS/) BENSON D R.
(LODE/) LODGES M J.
(MITC/) MITCHAM J L.
(KING/) KING G E.

Benson DR, Lodes MJ, Mitcham JL, King GE;

WPI: 2002-598720/64.

Composition for detecting and treating ovarian cancer, comprises a specific polypeptide, polynucleotide, T cell population, or antigen presenting cell -

Example 1; Page 113; 188bp; English.

The present invention relates to a method of detecting the presence of ovarian cancer in a patient, involving detecting ovarian cancer associated polynucleotides. The method is not only used to detect the presence of cancer, preferably ovarian cancer in a patient, but also is used to stimulate and/or expand T cells specific for an ovarian tumour protein. The sequences can be used in vaccines used to treat cancer. The present sequence is an ovarian cancer associated coding sequence.

Sequence 436 BP; 89 A; 102 C; 98 G; 147 T; 0 other;

Cy Match 39.6%; Score 270; DB 24; Length 436;
Local Similarity 100.0%; Pred. No. 5.2e-71;
Indels 0; Gaps 0;

413 AGACTTTCATGAGAGCTCACTGTTGGCTGCAAGCACTGACCACTGATACCA 472
1 AGACTTTCATGAGAGCTCACTGTTGGCTGCAAGCACTGACCACTGATACCA 60
473 CCATACCTGAGAGAGCTGTTGCTCCCTGAGGCGGCAATACCACTGATACCA 532
61 CCATACCTGAGAGAGCTGTTGCTCCCTGAGGCGGCAATACCACTGATACCA 120
533 AAGATTGTCATGAGAGAGCTGTTGCTCCCTGAGGCGGCAATACCACTGATACCA 592
121 AAGATTGTCATGAGAGAGCTGTTGCTCCCTGAGGCGGCAATACCACTGATACCA 180
593 CAGCATTTGGTGTAGCTGCTATTAATGATCTTGAAGATTTCTGAGCATGCTGCT 652
181 CAGCATTTGGTGTAGCTGCTATTAATGATCTTGAAGATTTCTGAGCATGCTGCT 240
653 GTGGCATCCGGAAGAGCTCGGTGATGAG 682
241 GTGGCATCCGGAAGAGCTCGGTGATGAG 270

RESULT 6
ABX73019
ID ABX73019 standard; cDNA; 436 BP.
XX
AC ABX73019;
XX
DT 14-MAR-2003 (first entry)
XX
DR Human ovarian carcinoma antigen partial cDNA sequence #285.
XX
KW Human; cancer detection; ovarian carcinoma antigen; ovarian cancer;
XX tumour antigen; tumour; OV2; OV3; OV6; OV9; OV10; OV12; OV14; OV17;
XX OV18; OV23; OV24; OV27; OV41; OV54; OV57; gene; ss.
XX
OS Homo sapiens.
XX
PN US6468758-B1.
XX
PD 22-OCT-2002.
XX
PF 16-SEP-1999; 99US-0397787.
XX
PR 23-SEP-1998; 98US-0159320.
XX 08-FEB-1999; 99US-0246429.
XX
PA (CORI-) CORIXA CORP.
XX
PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX
DR WPI: 2003-147101/14.
XX

Determining presence or absence of cancer in patient by contacting patient sample with oligonucleotide that hybridises to polynucleotide encoding ovarian carcinoma antigen, and detecting amount of hybridisation in sample -

Example 1; Column 217-218; 152bp; English.

The present invention relates to compositions and methods for determining presence or absence of cancer in patient. The method comprising contacting a biological sample with an oligonucleotide that hybridises to partial polynucleotide sequence encoding for human ovarian carcinoma antigen, or its complement. The method and compositions are useful for the therapy and diagnosis of ovarian cancer. The method can be used to identify tumour antigens that are secreted from ovarian carcinoma and/or other tumours. Effective cancer detection is achieved using the method of the invention.

Sequence 436 BP; 89 A; 102 C; 98 G; 147 T; 0 other;

Query Match 39.6%; Score 270; DB 25; Length 436;
Best Local Similarity 100.0%; Pred. No. 5.2e-71;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

413 AGACTTTCATGAGAGCTCACTGTTGGCTGCAAGCACTGACCACTGATACCA 472
1 AGACTTTCATGAGAGCTCACTGTTGGCTGCAAGCACTGACCACTGATACCA 60
473 CCATACCTGAGAGAGCTGTTGCTCCCTGAGGCGGCAATACCACTGATACCA 532
61 CCATACCTGAGAGAGCTGTTGCTCCCTGAGGCGGCAATACCACTGATACCA 120
533 AAGATTGTCATGAGAGAGCTGTTGCTCCCTGAGGCGGCAATACCACTGATACCA 592
121 AAGATTGTCATGAGAGAGCTGTTGCTCCCTGAGGCGGCAATACCACTGATACCA 180
593 CAGCATTTGGTGTAGCTGCTATTAATGATCTTGAAGATTTCTGAGCATGCTGCT 652
181 CAGCATTTGGTGTAGCTGCTATTAATGATCTTGAAGATTTCTGAGCATGCTGCT 240

653 GTGGCATCCGGAACAGCTCCGTTACTGAG 682
 |||||
 241 GTGGCATCCGGAACAGCTCCGTTACTGAG 270

14/c
 AH83814 standard; cDNA; 429 BP.
 AH83814;

5-SBP-2001 (first entry)

uman ovarian tumour associated polynucleotide sequence SEQ ID NO:1438.

uman; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
 immunogenic; vaccine; ss.

omo sapiens.

2200151513-A2.

9-JUL-2001.

6-JAN-2001; 2001MO-US01575.

4-JAN-2000; 2000US-0176722.

CORI-) CORIXA CORP.

lgate PA;

PI; 2001-425866/45.

ovel ovarian tumor proteins, and nucleic acids encoding them, used to
 treat and diagnose cancers, particularly ovarian cancer.

laim 5; Page 326; 338pp; English.

AH82377 to AAH8378 represent human ovarian tumour-associated
 polynucleotide sequences which encode ovarian tumour proteins. The
 ovarian tumour protein and polynucleotide sequences have cytostatic
 activity and can be used in gene therapy and vaccine production. The
 ovarian tumour proteins and polynucleotides can be used to inhibit
 the development of cancer, particularly ovarian cancer. They can also
 be used to diagnose the onset and progression of cancer.

sequence 429 BP; 94 A; 116 C; 126 G; 93 T; 0 other;

Match 32.1%; Score 219.2; DB 22; Length 429;
 Local Similarity 82.6%; Pred. No. 1.1e-55;

es 251; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

379 GATGATGATGCAACAAATGCAAGGCTGTGTAAGCTTTCATGAGAGCTCACTGT 438
 |||||

319 GGTGATGATGCAACAAATGCAAGGCTGTGTAAGCTTTCATGAGAGCTTGAAGTGC 260
 |||||

439 TGTGCTCCAAAGCACTGACCACTGATCACTACCACTGATGAGAAAGCTGTGCTCC 498
 |||||

259 TGTGCTCCAAAGCACTGACCACTGATCACTACCACTGATGAGAAAGCTGTGCTCC 200
 |||||

499 TCGAGCGGCAACATCACTACCCCTTACTGAGCAAGATGTCATCAAGAAATGATGAG 558
 |||||

199 TCGAGCGGCAACATCACTACCCCTTACTGAGCAAGATGTCATCAAGAAATGATGAG 140
 |||||

559 CTCTTCTCTGGGAAGCTGATCTCAATGAAATGCAAGCAATGAGCTGCTCATTTAG 618
 |||||

139 CTCTTCTCTGGGAAGCTGATCTCAATGAAATGCAAGCAATGAGCTGCTCATTTAG 80
 |||||

619 ATCTTTGATGATGATTTGAGCATGATGCTGTGTCATGAGCAATCCGAAACAGCTCCGTGAC 678
 |||||
 79 ATCTTTGATGATGATTTGAGCATGATGCTGTGTCATGAGCAATCCGAAACAGCTCCGTGAC 20
 |||||

OY 679 TGAG 682
 |||||
 Db 19 TGAG 16

RESULT 8
 AAS25168/c
 ID AAS25168 standard; cDNA; 316 BP.
 XX
 AC AAS25168;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human ovarian PCR-subtracted cDNA library clone #1349.
 XX
 KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
 KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
 KW primer; probe.
 XX
 OS Homo sapiens.
 XX
 PN W0200157207-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001MO-US03733.
 XX
 PR 04-FEB-2000; 2000US-0180403.
 PR 28-MAR-2000; 2000US-0192745.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Algate PA, Mannion J;
 XX
 DR WPI; 2001-488979/53.
 XX
 PT New polynucleotides encoding ovarian tumour proteins, useful for
 PT treating ovarian cancer, and as probes, primers, and markers of cancer
 PT progression
 XX
 PS Example 1; page 319; 378pp; English.
 XX
 XX The invention comprises compositions used for the therapy and diagnosis
 CC of ovarian cancer. The compositions comprise one or more ovarian tumour
 CC proteins, their associated polynucleotides, or immunogenic portions of
 CC the proteins. The ovarian tumour polynucleotides and immunogenic portions
 CC are useful for stimulating and/or expanding T cells specific for a tumour
 CC protein. They are also useful for inhibiting the development of cancer in
 CC a patient with an ovarian tumour DNA or protein by incubating isolated
 CC T-cells allowing them to proliferate, and administering to the patient.
 CC The sequences can be used as markers for cancer, for example, to monitor
 CC ovarian cancer progression. Probes and primers are useful in nucleic acid
 CC hybridisation, in detecting the presence of complementary sequences in a
 CC given sample, for preparing mutant species and for preparing other
 CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25238-AAS25549
 CC represent human ovarian tumour protein cDNA clones.
 XX
 SO Sequence 316 BP; 71 A; 83 C; 94 G; 68 T; 0 other;

Query Match 31.9%; Score 217.8; DB 22; Length 316;
 Best Local Similarity 82.7%; Pred. No. 2.4e-55;

Matches 249; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

OY 382 GATGATGCAACAAATGCAAGGCTGTGTAAGCTTTCATGAGAGCTCACTGTGT 441
 |||||

Db 316 GATGATGCAACAAATGCAAGGCTGTGTAAGCTTTCATGAGAGCTTGAAGTGC 257
 |||||

OY 442 GGCTCCAGCACTGACCACTGATCACTACCACTGATGAGCAAGCTGTGCTCCCTCA 501
 |||||

Db 256 GGCTCCAGCACTGACCTGTTGACCACTGATGCTCAAGAAATGTCCTCTG 197
 |||||
 OY 502 GGCGGCAATCACTACCCCTTACTGAGCAAGATGTCATCAAGAAATGATGAGCTC 561
 |||||

196 GGCAGCAACATCATCTGACCACTCTTCAGAGAGACTGCCAGAGATCGATGACCTC 137
 562 TTCTCGGAGAGCTGACCTGATTTGAAATTCAGCCATTTGTGAGCTGATTTATGATC 621
 136 TTCTCGGAGAGCTGACCTGATTTGAAATTCAGCCATTTGTGAGCTGATTTATGATC 77
 622 TTGAGATGATTTGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 681
 76 TTGAGATGATTTGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 17
 682 G 682
 16 G 16

T 9
 432/c
 AAF68432 standard; cDNA; 311 BP.
 AAF68432;

12-APR-2001 (first entry)
 Human lung tumour protein related nucleotide sequence SEQ ID NO:367.

Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
 lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
 cytostatic; antisense inhibition; ss.

Homo sapiens.
 MO200100828-A2.

04-JAN-2001.
 30-JUN-2000; 2000MO-US18061.

30-JUN-1999; 99US-0346492.
 15-OCT-1999; 99US-0419356.
 17-DEC-1999; 99US-0468867.
 30-DEC-1999; 99US-0476300.
 06-MAR-2000; 2000US-0519642.
 22-MAR-2000; 2000US-0531077.
 10-APR-2000; 2000US-0546259.
 27-APR-2000; 2000US-0560406.
 05-JUN-2000; 2000US-0589184.

(CORI-) CORIXA CORP.
 Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
 Retter MW, Mannion J;
 4PI; 2001-071488/08.

lung tumor-associated proteins and the nucleic acids that encode them,
 useful for preventing, diagnosing and treating lung cancer -
 Claim 4; Page 280; 436pp; English.

The present invention describes immunogenic portions of lung tumour-
 associated proteins (I) and the nucleic acids (NAs) that encode them.
 (I) have cytostatic activity and can be used in gene therapy, antisense
 inhibition and in vaccines. The NAs and the lung tumour-associated
 proteins they encode may be used in the prevention, treatment and
 diagnosis of diseases associated with their inappropriate expression,
 especially lung cancers. For example, the NAs may be administered to
 treat diseases by rectifying mutations or deletions in a patient's genome
 that affect the activity of the protein by expressing inactive proteins
 or to supplement the patient's own production of (I). Additionally, the
 NAs may be used to produce the lung tumour associated protein, according
 to standard recombinant DNA methodology. Conversely, antisense NA
 molecules may be administered to down regulate protein expression by

CC binding with the cells own genes and preventing their expression. The NA
 CC and complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar NA sequences in
 CC samples, and hence which patients may be in need of treatment for lung
 CC cancer. The (I) may be used as antigens in the production of antibodies
 CC and in assays to identify modulators (agonists and antagonists) of the
 CC expression and activity of the protein. AAF68083 to AAF68878 and
 CC AAF68848 to AAF68878 represent human lung tumour protein related
 CC nucleotide and protein sequences which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 311 BP; 70 A; 81 C; 94 G; 66 T; 0 other;

Query Match 31.3%; Score 213.4; DB 22; Length 311;
 Best Local Similarity 82.7%; Pred. No. 5e-54;
 Matches 244; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 388 GCCAACAATGCCAAGCTGTGTGAAGACTTTCATGAGACCTCACTGTTGTGCTCC 447
 DB 310 GCCAACAATGCCAAGCTGTGTGAAGACTTTCATGAGACCTCACTGTTGTGCTCC 251
 QY 448 AAGCACTGACCACTGACCTACCACTGACCTGACCTGACCTGACCTGACCTGAC 507
 DB 250 AAGCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 191
 QY 508 AACATCTACCCCTTACTGACGAGCAAGATTTCATGAGAAATGATGAGCTTCTCT 567
 DB 190 AACATCTACCAACCTCTTCAAGGAGCACTGACCAAGATGATGAGCTTCTCTC 131
 QY 568 GGGAGCTGTACCTGATGAGAAATTCAGCCATTTGTGAGCTGATTTATGATTTGAG 627
 DB 130 GGGAGCTGTACCTGATGAGAAATTCAGCCATTTGTGAGCTGATTTATGATTTGAG 71
 QY 628 ATGATCTGACCATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 682
 DB 70 ATGATCTGACCATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 16

RESULT 10
 ABK38343/c
 ID ABK38343 standard; cDNA; 311 BP.
 AC ABK38343;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE cDNA encoding clone #26488 of lung tumour protein.
 XX
 KW Lung tumour; cancer; T cell; immune response stimulator;
 KW cytostatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200204514-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 10-JUL-2001; 2001WO-US22058.
 XX
 PR 11-JUN-2000; 2000US-0614124.
 PR 29-AUG-2000; 2000US-0615163.
 PR 08-SEP-2000; 2000US-0658824.
 PR 26-SEP-2000; 2000US-0671325.
 PR 06-OCT-2000; 2000US-0677419.
 PR 30-OCT-2000; 2000US-0702705.
 PR 13-DEC-2000; 2000US-0736457.
 PR 03-MAY-2001; 2001US-0849626.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Wang T, Matenabe Y, Henderson RA, Johnson JC, Retter MW,
 PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
 PI Mcnabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;

PI: 2002-164634/21.

ovel polynucleotide encoding a lung tumour polypeptide useful for stimulating and/or expanding T cells specific for a tumour protein - sample 1; SEQ ID No 367; 223bp; English.

he invention describes an isolated polynucleotide and polypeptide useful for stimulating and/or expanding T cells specific for a tumour protein for determining the presence of a cancer in a patient. A composition containing the polynucleotide and/or polypeptide is useful for treating a lung cancer in a patient. The polypeptide is useful for moving tumour cells from a biological sample. The polynucleotide is also useful as probe or primer to detect the level of mRNA encoding a tumour protein. This sequence encodes a lung tumour associated protein r protein fragment, described in the method of the invention.

te: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences.

sequence 311 BP; 70 A; 81 C; 94 G; 66 T; 0 other;

y Match 31.3%; Score 213.4; DB 24; Length 311; Local Similarity 82.7%; Pred. No. 5e-54; Mismatches 51; Indels 0; Gaps 0;

388 GCCAACAAATGCCAAGCTGTGTGAGAACTTTCCATAGACGCTCAACTGTGTGCTCC 447
310 GCCAACAAATGCCAAGCTGTGTGAGAACTTTCCATAGACGCTCAACTGTGTGCTCC 251
448 AACGCACTGACCACTGACTATACCACTATGAGAAAGCCTGTGTGCTCAAGGCGC 507
250 AGCAGCTGACTGTGTGAGCACTGCTGAGCAAGCAATTTGTCTCTGGGCGAC 191
508 AACGCACTGACCACTGACTATACCACTATGAGAAAGCCTGTGTGCTCAAGGCGC 567
190 AACATCTGACCACTGACTATACCACTATGAGAAAGCCTGTGTGCTCAAGGCGC 131
568 GGGAGGCTGACCTGACTATGAGCACTGCTGAGCAAGCAATTTGTCTCTGGGCGC 627
130 GGGAGGCTGACCTGACTATGAGCACTGCTGAGCAAGCAATTTGTCTCTGGGCGC 71
628 ATGATCTGACCACTGACTATACCACTATGAGAAAGCCTGTGTGCTCAAGGCGC 682
70 ATGATCTGACCACTGACTATGAGCACTGCTGAGCAAGCAATTTGTCTCTGGGCGC 16

11 72/c
2A10672 standard; cDNA; 311 BP.
2A10672:

5-JUN-2003 (first entry)

uman lung neuroendocrine carcinoma library MLN1 cDNA, SEQ ID 367.

uman lung cancer; ss; lung tumour; cytostatic; vaccine;
cell expansion; CD4; CD8.

omo sapiens.

32002197669-A1.

5-DEC-2002.

5-MAY-2001; 2001US-0849626.

5-DEC-2000; 2000US-0736457.

BANGUR C S.

PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.

PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
XX WPI; 2003-352750/33.

Novel lung cancer polynucleotide encoding lung cancer protein, useful for detecting the presence of lung cancer in a patient, and in pharmaceutical compositions, e.g. vaccines, for treating lung cancer - Example 1; Page -; 72pp; English.

CC The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences
CC mentioned in the specification, or a sequence (S2) mentioned in
CC specification, complement of S1, sequences consisting of at least 20
CC contiguous residues of S1, sequences that hybridise to S1, sequences
CC having 75%, preferably 90%, identity to S1, or degenerate variants of
CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)
CC selected from any one of the 4 amino acid sequences mentioned in the
CC specification, a sequence encoded by the polynucleotide, or sequences
CC having at least 70%, preferably 90%, identity to a sequence encoded by
CC the polynucleotide), an expression vector comprising the polynucleotide
CC operably linked to an expression control sequence, a host cell
CC transformed or transfected with the vector, an isolated antibody (or its
CC antigen-binding fragment) that specifically binds to the polypeptide,
CC detecting the presence of a cancer in a patient, a fusion protein
CC comprising the polypeptide, an oligonucleotide that hybridises to
CC S1 under moderately stringent conditions, stimulating and/or expanding T
CC cells specific for a tumour protein (comprising contacting T cells with
CC the polynucleotide, protein or antigen-presenting cells, under conditions
CC and for a time sufficient to permit the stimulation and/or expansion of T
CC cells) and inhibiting the development of a cancer in a patient (by the
CC incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with the
CC polynucleotide, protein or antigen presenting cells that express the
CC polynucleotide, such that T cells proliferate, administering to the
CC patient an effective amount of the proliferated T cells, and thus
CC inhibiting the development of a cancer in the patient. The
CC polynucleotide, protein and cells are useful in a composition for
CC stimulating an immune response in a patient, and for treating a cancer in
CC a patient (particularly lung cancer). The oligonucleotide is useful for
CC determining the presence of a cancer in a patient. The protein and
CC oligonucleotides are useful in pharmaceutical compositions, e.g.
CC vaccines. The polynucleotide is also useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. An amplified portion of the polynucleotide is
CC useful for isolating a full-length gene from a suitable library.
CC The present sequence is a cDNA (full length, extended or partial)
CC isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from the USPTO
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.

XX Sequence 311 BP; 70 A; 81 C; 94 G; 66 T; 0 other;
Query Match 31.3%; Score 213.4; DB 25; Length 311;
Best Local Similarity 82.7%; Pred. No. 5e-54;
Matches 244; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

OY 388 GCCAACAAATGCCAAGCTGTGTGAGAACTTTCCATAGACGCTCAACTGTGTGCTCC 447
DB 310 GCCAACAAATGCCAAGCTGTGTGAGAACTTTCCATAGACGCTCAACTGTGTGCTCC 251
OY 448 AACGCACTGACCACTGACTATACCACTATGAGAAAGCCTGTGTGCTCAAGGCGC 507

250 AGCACTGACCTGCTTGGACCACTGAGTGTCAAGAAATTGTCCTCGGGGAGC 191
 508 AACATCTGACCCCTTACTGAGCAAGATTGTGATGAGAAATGAGTCTTCT 567
 190 AACATCTGACCAACCTTTCAGAGAGAGAGCAACCAAGATGATGACCTTCTCC 131
 568 GGGAGCTGTACTCTGATTTGATGAGCAATGTTGGTGTGATGATGATCTTGG 627
 130 GGGAGCTGTACTCTGATTTGATGAGCAATGTTGGTGTGATGATGATCTTGG 71
 628 ATGATTTGAGCATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 682
 70 ATGATTTGAGCATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 16
 '623/C
 AEX99623 standard; cDNA; 311 BP.
 AEX99623;
 22-MAY-2003 (first entry)
 Lung cancer therapy and diagnosis associated cDNA #350.
 Lung cancer; cytostatic; vaccine; gene therapy; cancer;
 gene; ss.
 Homo sapiens.
 US2002172952-A1.
 21-NOV-2002.
 10-JUL-2001; 2001US-0902941.
 30-JUN-1999; 99US-0346492.
 15-OCT-1999; 99US-0419356.
 17-DEC-1999; 99US-0466867.
 30-DEC-1999; 99US-0476300.
 06-MAR-2000; 2000US-0519642.
 22-MAR-2000; 2000US-0533077.
 10-APR-2000; 2000US-0546259.
 27-APR-2000; 2000US-0560406.
 05-JUN-2000; 2000US-0589184.
 11-JUL-2000; 2000US-0614124.
 29-AUG-2000; 2000US-0651563.
 08-SEP-2000; 2000US-0658824.
 26-SEP-2000; 2000US-0671325.
 06-OCT-2000; 2000US-0677419.
 30-OCT-2000; 2000US-0702705.
 13-DEC-2000; 2000US-0736457.
 23-MAY-2001; 2001US-0849626.
 (CORI-) CORIXA CORP.
 Henderson RA, Wang T, Matanabe Y, Johnson JC, Retter MW, Durham M;
 Carter D, Panger GR, Vedvick TS, Bangur CS, McNabb A;
 WPI; 2003-328427/31.
 few polynucleotide, useful for preparing a composition for treating or
 inhibiting development of cancer, e.g. lung cancer
 Example 1; SEQ ID NO 367; 82bp; English.
 The invention describes an isolated polynucleotide comprising one of 32
 sequences, complement or degenerate variants of them. The polynucleotide
 is useful for preparing a composition e.g. a vaccine or for gene therapy,
 for treating or inhibiting development of cancer, e.g. lung cancer.
 This sequence represents a polynucleotide associated with the
 compositions and methods for the therapy and diagnosis of lung cancer.

50 Sequence 311 BP; 70 A; 81 C; 94 G; 66 T; 0 other;
 Query Match 31.3%; Score 213.4; DB 25; Length 311;
 Best Local Similarity 82.7%; Pred. No. 5e-54;
 Matches 244; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 388 GCCAACAATCCCAAGCTGTGTGAGAGACTTTTCATGAGAGCTGATGCTGCTCC 447
 DB 310 GCCAACAAGCCCAAGGCTGTGTGAGAGACTTTTCATGAGAGCTGATGCTGCTCC 251
 QY 448 AACGACCTGACCACTGTACTGACCACTGACCACTGACCACTGACCACTGACCA 507
 DB 250 AGCACTGACCTGCTTGTGACCACTGACCACTGACCACTGACCACTGACCACTG 191
 QY 508 AACATCTGACCCCTTACTGAGCAAGATTGTGATGAGAAATGAGTCTTCTCT 567
 DB 190 AACATCTGACCAACCTTTCAGAGAGAGAGCAACCAAGATGATGACCTTCTCC 131
 QY 568 GGGAGCTGTACTCTGATTTGATGAGCAATGTTGGTGTGATGATGATCTTGG 627
 DB 130 GGGAGCTGTACTCTGATTTGATGAGCAATGTTGGTGTGATGATGATCTTGG 71
 QY 628 ATGATTTGAGCATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 682
 DB 70 ATGATTTGAGCATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 16
 RESULT 13
 AAX37230
 ID AAX37230 standard; DNA; 357 BP.
 AC AAX37230;
 XX
 DT 07-JUL-1999 (first entry)
 DE DNA sequence of EC2 fragment cloned in pThio-His C.
 XX
 KW CD81 protein; receptor; Hepatitis C virus; HCV; E2 protein; infection;
 KM immunogen; therapy; diagnosis; EC2 region; ss.
 OS Homo sapiens.
 XX
 PN WO9918198-A1.
 XX
 PD 15-APR-1999.
 XX
 PF 06-OCT-1998; 98WO-1B01628.
 XX
 PR 23-JUN-1998; 98GB-0013560.
 XX
 PR 06-OCT-1997; 97GB-0021182.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Abirignani S, Grandi G;
 XX
 DR WPI; 1999-264018/22.
 XX
 DR P-PSDB; AAY09154.
 XX
 PT Hepatitis C receptor protein CD81
 XX
 PS Example 6; Fig 5; 55bp; English.
 PS
 CC The invention relates to the use of the CD81 protein that is identified
 CC as a receptor for Hepatitis C virus (HCV). Compounds that specifically
 CC bind to CD81 can be used for therapy and diagnosis of HCV. The CD81
 CC protein binds to HCV, preferably to the E2 protein. The CD81 protein can
 CC be used to raise antibodies, which can then be used in the diagnosis or
 CC therapy of HCV. The antibodies can be used to prevent the virus binding
 CC to patient cells and being internalized. The CD81 protein can also be
 CC used to treat HCV infection. The CD81 protein, a functional equivalent,
 CC or a compound that binds to CD81, can be used in the manufacture of a
 CC medicament for the treatment and diagnosis of a HCV infection. The CD81
 CC protein can also be used as a protective immunogen in the control of

CV. The present sequence represents a DNA coding a EC2 fragment. EC2 is the region in CD81 that can bind to the E2 protein of HCV.

sequence 357 BP; 87 A; 101 C; 96 G; 73 T; 0 other;

Y Match 28.6%; Score 195.2; DB 20; Length 357;

Local Similarity 81.1%; Pred. No. 1.6e-48; Indels 0; Gaps 0;

hes 227; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

298 ATCTGGGGCTTGTAAACCAAGACGATCGCAAGATGTGAAGACGTTCTATGACGAG 357

70 AGCTGAGAGCTTGTCAACAGACGATCGCAAGATGTGAAGACGTTCTATGACGAG 129

358 GCCCTTACGACGCTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 417

130 GCCCTTACGACGCTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 189

418 TTCCATGAGAGCTTCAACGCTTGTGCTTCAACGCTTGTGCTTCAACGCTTGTGCTT 477

190 TTCCATGAGAGCTTCAACGCTTGTGCTTCAACGCTTGTGCTTCAACGCTTGTGCTT 249

478 CTGAGACAGCCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 537

250 CTGAGACAGCCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 309

538 TGTGATCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 577

310 TGCCACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 349

14

31 AX37231 standard; DNA; 348 BP.

AX37231;

7-JUL-1999 (first entry)

NA sequence of EC2-His6 fragment cloned into pGEX-KG.

DB1 protein; receptor; Hepatitis C virus; HCV; E2 protein; infection;

immunogen; therapy; diagnosis; EC2 region; ss.

omo sapiens.

09918198-A1.

5-APR-1999.

6-OCT-1998; 98WO-IB01628.

3-JUN-1998; 98GB-0013560.

6-OCT-1997; 97GB-0021182.

CHIR-) CHIRON SPA.

brlgnani S, Grandi G;

PI; 1999-264018/22.

-PSDB; AAY09155.

epatitis C receptor protein CD81

xample 6; Fig 8; 55pp; English.

he invention relates to the use of the CD81 protein that is identified as a receptor for Hepatitis C virus (HCV). Compounds that specifically bind to CD81 can be used for therapy and diagnosis of HCV. The CD81 protein binds to HCV, preferably to the E2 protein. The CD81 protein can be used to raise antibodies, which can then be used in the diagnosis or therapy of HCV. The antibodies can be used to prevent the virus binding to patient cells and being internalized. The CD81 protein can also be used to treat HCV infection. The CD81 protein, a functional equivalent,

CC or a compound that binds to CD81, can be used in the manufacture of a CC medicant for the treatment and diagnosis of a HCV infection. The CD81 protein can also be used as a protective immunogen in the control of CC HCV. The present sequence represents a DNA coding a EC2 fragment. EC2 is the region in CD81 that can bind to the E2 protein of HCV.

Sequence 348 BP; 84 A; 96 C; 90 G; 78 T; 0 other;

Query Match 28.0%; Score 190.8; DB 20; Length 348;

Best Local Similarity 81.0%; Pred. No. 3.4e-47;

Matches 222; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 307 TTGTTAAACCAAGACGATCGCAAGATGTGAAGACGTTCTATGACGAGCCCTTCA 366

Db 52 TTGTTAAACCAAGACGATCGCAAGATGTGAAGACGTTCTATGACGAGCCCTTCA 111

QY 367 CAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426

Db 112 CAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 171

QY 427 ACCCTCAACTGTTGT 486

Db 172 ACCCTCAACTGTTGT 231

QY 487 AGCTGTGTCTGTCAAGCGGCAACATCTCAACCTTCTCTCTCTCTCTCTCTCTCT 546

Db 232 AATTTGTGTCTGTCAAGCGGCAACATCTCAACCTTCTCTCTCTCTCTCTCTCTCT 291

QY 547 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 580

Db 292 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325

RESULT 15

AAQ29182

ID AAQ29182 standard; DNA; 1120 BP.

AC AAQ29182;

DT 25-MAR-2003 (updated)

DT 05-MAR-1993 (first entry)

XX DNA encoding metastasis controlling peptide.

XX M31-15 monoclonal antibody; cancer cell motility; prophylaxis;

XX treatment; lung; glioblastoma; ss.

XX Homo sapiens.

XX Key

XX CDS

XX mat_peptide

XX /tag= a

XX /tag= b

XX /note= "mature peptide"

XX BP508417-A2.

XX 14-OCT-1992.

XX 09-APR-1992; 92BP-0106093.

XX 12-APR-1991; 91JP-0079996.

XX 17-APR-1991; 91JP-0085396.

XX 07-FEB-1992; 92JP-0022321.

XX (TAKA) TAKADA CHEM IND LTD.

XX (MIVA) MIYAKE M.

XX Ikeyama S, Koyama M, Miyake M, Senoo M;

XX WPI; 1992-341723/42.

XX P-PSDB; AAR7525.

New monoclonal antibody M31-15 specific for metastasis-controlling peptide - useful for treating and preventing cancer and metastasis
Claim 8; Fig 4; 34pp; English.

The sequence is that of DNA encoding a cancer cell surface protein which is capable of suppressing the motility of cancer cells. It is specifically recognised by the monoclonal (Mab) antibody M31-15 which is useful for suppressing cancer metastasis. The polypeptide encoded by the DNA and Mab M31-15 are therefore useful in the prophylaxis and treatment of cancer, e.g. lung, glioblastoma and other metastatic cancers.
(Updated on 25-MAR-2003 to correct FN field.)

Sequence 1120 BP; 250 A; 256 C; 278 G; 336 T; 0 other;

100% Match 22.1%; Score 151; DB 13; Length 1120;
Local Similarity 55.3%; Pred. No. 6.1e-35;
Cons 373; Conservative 0; Mismatches 275; Indels 27; Gaps 3;

```
1 AAATACCTGCTCTGCTGCTCAATTTGCTCTTCTGCTGCTGAGCGCTGATCTTAGGT 60
  |||||
142 AAATACCTGCTGCTGCTGCTCAATTTGCTCTTCTGCTGCTGAGCGCTGATCTTAGGT 120
  |||||
61 GTAGCTCTGCTGCTGCTGCTCAATTTGCTCTTCTGCTGCTGAGCGCTGATCTTAGGT 120
  |||||
202 ATTGACTATATGCTGCTGCTGCTCAATTTGCTCTTCTGCTGCTGAGCGCTGATCTTAGGT 261
  |||||
121 AACAAACGGACCCCAACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
  |||||
262 AATTAATTAATTC-----CAGCTTCAACAGAGGTATTTCTGATCGAGCGCGGCC 315
  |||||
181 GATGATATGTTTGTAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
  |||||
316 CTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375
  |||||
241 CTGGGAGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
  |||||
376 CTGGGAGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 435
  |||||
301 TGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
  |||||
436 TGGGATATTTCCCAAGAGTGAAGTGAATTAAGAAATTCAGAGATTTTCAAGAGACCC 495
  |||||
361 CTTCACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
  |||||
496 TACAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
  |||||
421 CATGAGAGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
  |||||
550 CACTATGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
  |||||
481 AGGAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
  |||||
610 TGGCCCAAGAG-----GAGTACTGAAACCTTCAAGCTGAAAGTCTGT 654
  |||||
541 CATCAAAAAATGATGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
  |||||
655 CCTGATGCTCAAAAGAGCTCTTCAAGATTAATTCATCATCATCGGCGGAGTGGGCAATC 714
  |||||
601 GTGTAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
  |||||
715 GGCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774
  |||||
661 CGGAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
  |||||
775 CGCAGGAACCGGAG 789
```

1	270	39.6	436	4	US-09-397-787-303	Sequence 303, Appl
2	213.4	31.3	311	4	US-09-702-705-367	Sequence 367, Appl
3	213.4	31.3	311	4	US-09-736-457-167	Sequence 367, Appl
4	151	22.1	687	1	US-08-254-493-2	Sequence 2, Appl
5	151	22.1	687	1	US-08-408-2228-2	Sequence 2, Appl
6	151	22.1	687	1	US-08-408-2228-3	Sequence 3, Appl
7	149.4	21.9	1120	1	US-08-254-493-3	Sequence 3, Appl
8	104.4	15.3	679	4	US-09-040-984-14	Sequence 14, Appl
9	104.4	15.3	679	4	US-09-123-917-14	Sequence 14, Appl
10	104.4	15.3	679	4	US-05-480-884A-14	Sequence 14, Appl
11	104.4	15.3	679	4	US-05-542-615A-14	Sequence 14, Appl
12	104.4	15.3	679	4	US-09-606-421B-14	Sequence 14, Appl
13	104.4	15.3	679	4	US-08-430-225A-19	Sequence 19, Appl
14	104.4	15.3	679	4	US-08-705-771-8	Sequence 8, Appl
15	66.8	9.8	1624	3	US-08-807-044-4	Sequence 4, Appl
16	64	9.4	1452	2	PCT-US91-04986-1	Sequence 1, Appl
17	64	9.4	1452	2	US-09-482-273-41	Sequence 41, Appl
18	62.8	9.2	1932	4	US-09-020-956-17	Sequence 17, Appl
19	61	8.9	1151	2	US-09-030-607-17	Sequence 17, Appl
20	60.4	8.9	740	3	US-09-439-313-17	Sequence 17, Appl
21	60.4	8.9	740	4	US-09-352-616A-17	Sequence 17, Appl
22	60.4	8.9	740	4	US-09-232-149A-17	Sequence 17, Appl
23	60.4	8.9	933	3	US-08-808-148-2	Sequence 2, Appl
24	60.4	8.9	1283	3	US-09-020-556-111	Sequence 111, Appl
25	60.4	8.9	1283	3	US-09-030-607-111	Sequence 111, Appl

28	60.4	8.9	1289	4	US-09-439-313-111	Sequence 111, Appl
29	60.4	8.9	1289	4	US-09-352-616A-111	Sequence 111, Appl
30	60.4	8.9	1289	4	US-09-232-148A-111	Sequence 111, Appl
31	60.2	8.8	1297	4	US-09-308-98A-1	Sequence 1, Appl
32	60	8.8	827	3	US-09-333-599-5	Sequence 5, Appl
33	60	8.8	827	4	US-09-499-781-5	Sequence 5, Appl
34	60	8.8	870	3	US-09-333-599-1	Sequence 1, Appl
35	60	8.8	870	4	US-09-429-781-1	Sequence 1, Appl
36	57.8	8.5	560	3	US-09-221-298-12	Sequence 12, Appl
37	56.8	8.3	729	3	US-09-020-955-13	Sequence 13, Appl
38	56.8	8.3	729	3	US-09-030-607-13	Sequence 13, Appl
39	56.8	8.3	729	4	US-09-439-313-13	Sequence 13, Appl
40	56.8	8.3	729	4	US-09-352-616A-13	Sequence 13, Appl
41	56.8	8.3	729	4	US-09-232-148A-13	Sequence 13, Appl
42	56.8	8.3	977	2	US-08-855-140-2	Sequence 2, Appl
43	56.8	8.3	977	4	US-09-015-433-938	Sequence 938, Appl
44	54.8	8.0	801	3	US-09-020-956-1	Sequence 16, Appl
45	54.8	8.0	801	3	US-09-030-607-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

```

1 Sequence303, Application US/09397787
2 Patent No. 6468758
3
4 GENERAL INFORMATION:
5
6 APPLICANT: Benson, Darin R.
7 APPLICANT: Kodes, Michael J.
8 APPLICANT: Mitcham, Jennifer L.
9 APPLICANT: King, Gordon E.
10 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
11 TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
12 FILE REFERENCE: 210121.466C
13 CURRENT APPLICATION NUMBER: US/09/397,787
14 CURRENT FILING DATE: 1999-09-16
15
16 NUMBER OF SEQ ID NOS: 334
17 SOFTWARE: FastSeq for Windows Version 3.0
18 SEQ ID NO: 303
19 LENGTH: 436
20 TYPE: DNA
21
22 ORGANISM: Homo sapien
23
24 US-09-397-787-303

```

Query Match	Score	DB 4	Length
39.6%	270	436	
100.0%	270	436	

Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	411	AGACCTTTCATGAGAGAGCTCAACGTTTGTGCTCCAAAGCATACCAACAACA	472
Db	1	AGACCTTTCATGAGAGAGCTCAACGTTTGTGCTCCAAAGCATACCAACA	60
Oy	473	CCATACCTGAGGAACAGCCTGTGTCCCTCAGAGCGGCAACATCTCACCCCTTAC	532
Db	61	CCATACCTGAGGAACAGCCTGTGTCCCTCAGAGCGGCAACATCTCACCCCTTAC	120
Oy	533	AGAGTTGTTCATGAGAAATTCAGTAGCTCTTCTGTGGAGAGCTGTACTCATTTG	592
Db	121	AGAGTTGTTCATGAGAAATTCAGTAGCTCTTCTGTGGAGAGCTGTACTCATTTG	180
Oy	593	CAGCCATTTGTGTACTCTGTCAATTATGATCTTTTGAGATGATTTCTGAGATG	652
Db	181	CAGCCATTTGTGTACTCTGTCAATTATGATCTTTTGAGATGATTTCTGAGATG	240
Oy	653	GTGGCATCCGGAAAGAGCTCCGTTACTGAG	682
Db	241	GTGGCATCCGGAAAGAGCTCCGTTACTGAG	270

RESULT 2
 US-09-702-705-367/c
 Sequence 367, Apolification US/09702705

nt No. 6504010
RAL INFORMATION:
LICANT: Wang, Tonglong
LICANT: Bangur, Chaitanya S.
LICANT: Lodes, Michael A.
LICANT: Panger, Gary
LICANT: Vedvick, Tom
LICANT: Carter, Darick
LICANT: Retter, Marc
LICANT: Mannion, Jane
LICANT: Pan, Liqun
LE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
LE OF INVENTION: DIAGNOSIS OF LUNG CANCER
E REFERENCE: 210121.478C14
RENT APPLICATION NUMBER: US/09/702,705
RENT FILING DATE: 2000-10-30
BER OF SEQ ID NOS: 1833
TMARK: FASTSEQ for Windows Version 3.0
ID NO 367
NGTH: 311
PR: DNA
GANISM: Homo sapien
702-705-367

y Match 31.3% Score 213.4; DB 4; Length 311;
Local Similarity 82.7% Pred. No. 6.9e-55;
hes 244; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
388 GCCAATGCGCAAGCTGTGTGTAAGACTTTCATGAGAGCTCACTGTGTGCTCC 447
310 GCCAACAAAGCGCAAGCTGTGTGTAAGACTTTCATGAGAGCTCACTGTGTGCTCC 251
448 AACGACCTGACGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 507
250 AGCAGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 191
508 AACATCTACACCCCTTACTGAGAGCAAGATTGTCTATCGAAGATGATGATGATGATG 567
190 AACATCTACAGCACTCTTCAAGAGAGCTGCCAGAGAGATGATGATGATGATGATGATG 131
568 GGAAGCTGTACTTCTATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 627
130 GGAAGCTGTACTTCTATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 71
628 ATGATCTGAGACATGCTGT 682
70 ATGATCTGAGACATGCTGT 16

3
736-457-367/c
ence 367, Application US/09736457
nt No. 6509448
RAL INFORMATION:
LICANT: Wang, Tonglong
LICANT: Bangur, Chaitanya S.
LICANT: Lodes, Michael A.
LICANT: Panger, Gary
LICANT: Vedvick, Tom
LICANT: Carter, Darick
LICANT: Retter, Marc
LICANT: Mannion, Jane
LICANT: Pan, Liqun
LICANT: Wang, Aljun
LE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
LE OF INVENTION: DIAGNOSIS OF LUNG CANCER
E REFERENCE: 210121.478C15
RENT APPLICATION NUMBER: US/09/736,457
RENT FILING DATE: 2000-12-13
BER OF SEQ ID NOS: 1864
TMARK: FASTSEQ for Windows Version 3.0
ID NO 367
NGTH: 311

TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-367

Query Match 31.3% Score 213.4; DB 4; Length 311;
Best Local Similarity 82.7% Pred. No. 6.9e-55;
Matches 244; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 388 GCCAATGCGCAAGCTGTGTGTAAGACTTTCATGAGAGCTCACTGTGTGCTCC 447
Db 310 GCCAACAAAGCGCAAGCTGTGTGTAAGACTTTCATGAGAGCTCACTGTGTGCTCC 251
Qy 448 AACGACCTGACGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 507
Db 250 AGCAGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 191
Qy 508 AACATCTACACCCCTTACTGAGAGCAAGATTGTCTATCGAAGATGATGATGATGATG 567
Db 190 AACATCTACAGCACTCTTCAAGAGAGCTGCCAGAGAGATGATGATGATGATGATG 131
Qy 568 GGAAGCTGTACTTCTATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 627
Db 130 GGAAGCTGTACTTCTATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 71
Qy 628 ATGATCTGAGACATGCTGT 682
Db 70 ATGATCTGAGACATGCTGT 16

RESULT 4
US-08-254-493-2
Sequence 2, Application US/08254493
Patent No. 5439886
GENERAL INFORMATION:
APPLICANT: IKEYAMA, SHUICHI
APPLICANT: KOYAMA, MASARU
APPLICANT: MIYAKI, MASAYUKI
APPLICANT: SENOO, MASAHARU
TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/254,493
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/865552
FILING DATE: 09-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 079996-1991
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 085396-1991
FILING DATE: 17-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 022321-1992
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, DAVID S.
REGISTRATION NUMBER: 34235
REFERENCE/DOCKET NUMBER: 41777

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRB UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: HUMAN
CELL TYPE: BREAST CARCINOMA
CELL LINE: ZR-75-1
-254-493-2

Match 22.1%; Score 151; DB 1; Length 687;
Local Similarity 55.3%; Pred. No. 6.6e-36;
ches 373; Conservative 0; Mismatches 275; Indels 27; Gaps 3;

```
1 AATACCTGCTCTTCTGCTTCAATTTCTGCTGCTGCTGAGGCGTATCTAGGT 60
31 AATACCTGCTGTTGGGATTTAATTCTTCTGCTGCTGCGGATTTGCTGCTTGC 90
61 GAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
91 ATTGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 150
121 AACAAACCGGACCCGAAACCTTCTAGGCGGATCTAATCTCTGCTGCTGCTG 180
151 AATATATATTC-----GAGCTTCAACAGAGGCTATATTTCTGATGCGAAC 204
181 GTGATGATTTTGTAGGCTTCTGCGGCTCTGAGGCTATGAGGCGATCCAGT 240
205 CTCATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 264
241 CTGGGAGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
265 CTGGGAGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 324
301 TGGGCGCTTGTAAACAGACAGATCGGCAAGATGTAAGAGCTTCTATGACAG 360
325 TGGGATATTTCCCAAGAGATGAGGATTAAGAGATTCAGAGATTTTACAGAG 384
361 CTTCAGCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 420
385 TACACAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 438
421 CATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
439 CACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 498
481 AGGACAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
499 TCCCCAGAGG-----GAGCTTCAACAGAGCTTCAACAGAGCTTCAAC 543
541 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
544 CTGATGCTATCAAGAGGCTTTCAGCAATTAATTCACATCATGAGGCGAGT 603
601 GTGATGCTGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
604 GGCATGCTGCTGCTATGATGATGATGATGATGATGATGATGATGATGAT 663
661 CGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 675
664 CGCAGAGAGCGGAG 678
```

US-08-408-222B-2
Sequence 2, Application US/08408222B
Patent No. 576727
GENERAL INFORMATION:
APPLICANT: Ikeyama, Shuichi
APPLICANT: Koyama, Masaru
APPLICANT: Miyake, Masayuki
APPLICANT: Senoo, Masaharu
TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: Dike, Bronsteijn, Roberts & Cushman
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTSEU Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,222B
FILING DATE: 22-MAR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/254,493
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-079996-1991
FILING DATE: 12-APR-1991
APPLICATION NUMBER: JP-085396-1991
FILING DATE: 14-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-022321-1992
FILING DATE: 07-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 41777-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: breast carcinoma
CELL LINE: ZR-75-1
US-08-408-222B-2

Query Match 22.1%; Score 151; DB 1; Length 687;
Best Local Similarity 55.3%; Pred. No. 6.6e-36;
Matches 373; Conservative 0; Mismatches 275; Indels 27; Gaps 3;

```
Qy 1 AATACCTGCTCTTCTGCTTCAATTTCTGCTGCTGCTGAGGCGTATCTAGGT 60
Db 31 AATACCTGCTGTTGGGATTTAATTCTTCTGCTGCTGCGGATTTGCTGCTTGC 90
Oy 61 GAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 91 ATTGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 150
```


121 AACAAACCGGACCCCAACACCTTCTAGCTGGGCACTACATCTCTGCTGGAGCT 180
 151 AATAATATATTC-----CAGCTTACACAGAGATATATTTGATCGGAGCGCGGCC 204
 181 GTATATATTTTGAAGCTTCTTGGGCTGCTATGGGCAATCCAGAGTCCAGTCTTG 240
 205 CTCATATATGCTGGGCTTCTTGGGCTGCTGCGGGGCTGTCAGAGAGTCCAGTCTATG 264
 241 CTGGGAGCTTCTTCACTGCTGCTGTGATCCCTGTTTTCCTGTGAGGAGGCTGGAGGATC 300
 265 CTGGGAGCTTCTTCTGGCTTCTTGTGATATTCGCAATGAAATAGCTGCGGCATC 324
 301 TGGGAGCTTCTTAAACAAAGACAGATGCGCAAGATGTAAGAGCTTCTATGACAGGCC 360
 325 TGGGAGATATTCACCAAGAGATGAGGTATTAAGAGATCCAGAGATTTTACAGAGACC 384
 361 CTTCAGCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 385 TACACACAGCTGTAACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 438
 421 CATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 439 CACTATGCTGTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 498
 481 AGGAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 499 TCGCCCAAGAG-----GAGCTTACAAACCTTCAACCGTGAAGTCTCTGT 543
 541 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 544 CCGTATGCTGCAACAAAGAGCTTCTGCAATTAATTTCCATCAATCGGCGAGTGGCATC 603
 601 GTGGTACTGCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 604 GGCATTCGCGGTGCTGCAATTAATTTGGCATATCTTCAGTATGATCTGTGCTGTATC 663
 661 CGGAAACAGCTCCGTG 675
 664 CGGAGAACCGCGAG 678

108-222B-3
 ance 3, Application US/08408222B
 at No. 5776727

GENERAL INFORMATION:
 APPLICANT: Ikeyama, Shuichi
 APPLICANT: Koyama, Masaru
 APPLICANT: Miyake, Masayuki
 APPLICANT: Senoo, Masaharu
 TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
 NUMBER OF SEQUENCES: 7
 ADDRESS/ADDRESS:
 ADDRESS/ADDRESS: Di, Bronstein, Roberts & Cushman
 STREET: 130 Water Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/408,222B
 FILING DATE: 22-MAR-1995
 CLASSIFICATION: 536
 RIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/254,493
 FILING DATE: 06-JUN-1994
 RIOR APPLICATION DATA:
 APPLICATION NUMBER: JP-079996-1991

FILING DATE: 12-APR-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP-085396-1991
 FILING DATE: 14-APR-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP-022321-1992
 FILING DATE: 07-FEB-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Resnick, David S.
 REGISTRATION NUMBER: 34,235
 REFERENCE/DOCKET NUMBER: 41777-DIV
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-523-3400
 TELEFAX: 617-523-6440
 INFORMATION FOR SEQ. ID NO. 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1120 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA to genomic RNA
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 ORGANISM: human
 CELL TYPE: breast carcinoma
 CELL LINE: ZR-75-1
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 112...795
 OTHER INFORMATION: E Mat peptide
 US-08-408-222B-3

Query Match 22.1%; Score 151; DB 1; Length 1120;
 Best Local Similarity 55.3%; Pred. No. 8-4e-36;
 Matches 373; Conservative 0; Mismatches 275; Indels 27; Gaps 3;

1 AATAATATATTC-----CAGCTTACACAGAGATATATTTGATCGGAGCGCGGCC 204
 142 AATAATATATTC-----CAGCTTACACAGAGATATATTTGATCGGAGCGCGGCC 204
 61 GTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 202 ATTGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 261
 121 AACAAACCGGACCCCAACACCTTCTAGCTGGGCACTACATCTCTGCTGGAGCT 180
 262 AATAATATATTC-----CAGCTTACACAGAGATATATTTGATCGGAGCGCGGCC 315
 181 GTATATATTTTGAAGCTTCTTGGGCTGCTATGGGCAATCCAGAGTCCAGTCTTG 240
 316 CTCATATATGCTGGGCTTCTTGGGCTGCTGCGGGGCTGTCAGAGAGTCCAGTCTATG 375
 241 CTGGGAGCTTCTTCACTGCTGCTGTGATCCCTGTTTTCCTGTGAGGAGGCTGGAGGATC 300
 376 CTGGAGCTTCTTCTGGCTTCTTGTGATATTCGCAATGAAATAGCTGCGGCATC 435
 301 TGGGAGCTTCTTAAACAAAGACAGATGCGCAAGATGTAAGAGCTTCTATGACAGGCC 360
 436 TGGGAGATATTCACCAAGAGATGAGGTATTAAGAGATCCAGAGATTTTACAGAGACC 495
 361 CTTCAGCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 496 TACACACAGCTGTAACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
 421 CATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 550 CACTATGCTGTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
 481 AGGAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 610 TCGCCCAAGAG-----GAGCTTACAAACCTTCAACCGTGAAGTCTCTGT 654

ORGANISM: HUMAN

TITLE OF INVENTION: OF LUNG CANCER

RESPONSE ADDRESS:
ADDRESS: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,984
FILING DATE: 18-MAR-1998
CLASSIFICATION:
TORTNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.456
ELECTRONIC COMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-282-6031
TELEX:
FORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 679 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
040-984-14
Y Match 15.3%; Score 104.4; DB 3; Length 679;
Local Similarity 59.0%; Pred. No. 7.4e-22;
hes 181; Conservative 0; Mismatches 120; Indels 6; Gaps 1;
1 AAATCCGCTCTTGTCTTCAATTCGCTTCTGCTGCGAGGCGTATCTAGT 60
|||||
132 AAATACCTGCGCTGCGATTAATTCATCTTCTGCTGCGAGGCGTATCTAGT 191
61 GTAGCTCTGCTGCTGCTGCTATGATTCAGACAGACAGGCGTATCTAGT 120
|||||
192 ATTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 245
|||||
121 AACAAACCGGACCCAGACCTTCTTACGTTGCGATCTTACATTCGCTGCGAGT 180
|||||
246 ACTAATNATNATNATNATTCAGCTTCTTACAGAGGCTTATTCATTCGATGCGATCCGCGC 305
|||||
181 GTGATGATTTTGTAGCTTCTGCGGCTCTATGCGGCGATCCAGAGTCCAGTGTCTG 240
|||||
306 CTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 365
|||||
241 CTGAGGACGCTTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
|||||
366 CTGAGGACGCTTCTGCGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
|||||
301 TGGGGCT 307
|||||
426 TGGGGAT 432

PRIOR FILING DATE: 1998-03-18
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 679
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (129)
OTHER INFORMATION: where n is a, c, g or t
LOCATION: (68)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (83)
OTHER INFORMATION: where n is a, c, g or t
LOCATION: (87)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (94)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (104)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (117)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (142)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (145)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (151)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (187)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (201)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (211)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (226)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (229)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (239)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (241)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (245)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (252)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (255)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (259)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (303)

Query Match	15.3%	Score 104.4;	DB 4;	Length 679;
Best Local Similarity	59.0%;	Pred. No. 7.4e-22;		
Matches 181; Conservative	0;	Mismatches 120;	Indels 6;	Gaps 1

Db 426 TGGGGAT 432

US-09-643-597-14
; Sequence 14, Application US/09643597

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Fan, Liqun

APPLICANT: Bangur, Chaltanya S.

APPLICANT: Hosken, Nancy

APPLICANT: Li, Samuel X

APPLICANT: Wang, Aijun

APPLICANT: Skeelky, Yasir A.W.

APPLICANT: McNeill, Patricia D.

1. TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

1. TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

CURRENT APPLICATION NUMBER: US/09/643,597

CURRENT FILING DATE: 2000-08-21

NUMBER OF SEQ ID NOS: 369
SOFTWARE: FastSeq for Windows v 2.0.0

SOFTWARE: TABCSDQ FOR WINDOWS VERSION 3.0
SEQ ID NO 14

LENGTH: 679

TYPE: DNA
ORGANISM: Homo sapien

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: E - A T C or C

US-09-643-597-14

Query Match

Best Local Similarity	59.0%;	Pred. NO. 7.4e-22;
Matches 181;	Conservative	0;
Mismatches	120;	Indels 6;
		Gaps 1

```

1 AAATACCTGCTCTGCTCTTCAATTGCTTCTGAGCTGAGAGGCTGATCTAGT 60
132 AAATACCTGCTCTGCTCTTCAATTGCTTCTGAGCTGAGAGGCTGATCTAGT 191
61 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
192 ATTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 245
121 AACAAACCGGACCCCAACACCTTCTAAGTGGGCTGCTGCTGCTGCTGCTG 180
246 ACTAATATATATATATATATATATATATATATATATATATATATATATAT 305
181 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
306 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 365
241 CTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
366 CTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 425
301 TGGGGCT 307
426 TGGGGAT 432

```

11
480-884A-14
ence 14, Application US/09480884A
nt No. 5482597

RL INFORMATION:
LICANT: Wang, Tonglong
LICANT: Pan, Liqun
LICANT: Hosken, Nancy A.
LICANT: Kalos, Michael D.
LICANT: Panger, Gary R.
LICANT: Panger, Gary R.
LE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
LE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
B REFERENCE: 210121.455C6
RENT APPLICATION NUMBER: US/09/480,884A
RENT FILING DATE: 2001-08-27
BER OF SEQ ID NOS: 330
TWARE: FastSeq for Windows Version 3.0
ID NO 14
NGTH: 679
PR: DNA
GANISM: Homo sapien
ATURE:
ME/KEY: misc_feature
CATION: (1)...(679)
HER INFORMATION: n = A,T,C or G
480-884A-14

Y Match 15.3%; Score 104.4; DB 4; Length 679;
Local Similarity 59.0%; Pred. No. 7.4e-22;
hes 181; Conservative 0; Mismatches 120; Indels 6; Gaps 1;

```

1 AAATACCTGCTCTGCTCTTCAATTGCTTCTGAGCTGAGAGGCTGATCTAGT 60
132 AAATACCTGCTCTGCTCTTCAATTGCTTCTGAGCTGAGAGGCTGATCTAGT 191
61 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
192 ATTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 245
121 AACAAACCGGACCCCAACACCTTCTAAGTGGGCTGCTGCTGCTGCTGCTG 180
246 ACTAATATATATATATATATATATATATATATATATATATATATATATAT 305
181 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
306 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 365

```

```

Qy 241 CTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 366 CTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 425
Qy 301 TGGGGCT 307
Db 426 TGGGGAT 432

```

RESULT 12

US-09-542-615A-14
Sequence 14, Application US/09542615A
Patent No. 6518256
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Pan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy A.
APPLICANT: Panger, Gary R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C8
CURRENT FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 350
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 679
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(679)
OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-14

Query Match

Best Local Similarity 15.3%; Score 104.4; DB 4; Length 679;
Best Local Similarity 59.0%; Pred. No. 7.4e-22;
Matches 181; Conservative 0; Mismatches 120; Indels 6; Gaps 1;

```

Qy 1 AAATACCTGCTCTGCTCTTCAATTGCTTCTGAGCTGAGAGGCTGATCTAGT 60
Db 132 AAATACCTGCTCTGCTCTTCAATTGCTTCTGAGCTGAGAGGCTGATCTAGT 191
Qy 61 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 192 ATTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 245
Qy 121 AACAAACCGGACCCCAACACCTTCTAAGTGGGCTGCTGCTGCTGCTGCTG 180
Db 246 ACTAATATATATATATATATATATATATATATATATATATATATATATAT 305
Qy 181 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
Db 306 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 365
Qy 241 CTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 366 CTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 425
Qy 301 TGGGGCT 307
Db 426 TGGGGAT 432

```

RESULT 13

US-09-606-421B-14
Sequence 14, Application US/09606421B
Patent No. 6531315
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Pan, Liqun

Page 9

```

/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERFECT 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/430.225A
/ FILING DATE: 28-APR-1995
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: RICHARD W. BORK
/ REGISTRATION NUMBER: 36,459
/ REFERENCE/DOCKET NUMBER: 2026-4172
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ TELEX: 421792
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1624 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
US-08-430-225A-19

Query Match          9.8%; Score 67; DB 3; Length 1624;
Best Local Similarity 52.8%; Pred. No. 2.2e-10;
Matches 179; Conservative 0; Mismatches 145; Indels 15; Gaps 14;

Db      1 A A A T A C C G C T C T C G T C T T C A A T T T G T T T C T G G C T G G C T G A G G G T A T C C T A G T 60
193 A A A T A C T T C T C T C C T C T T C A A C T T A T C T T T A T C C T G G C G C A G T A T C C T G G C 252

Qy      61 G T A C C T G T G G T T G C G T C A T G A T C C A C A G C A C A G C C T G T A C T T G A A C T G G A 120
253 T T C G G G G T G T G A T C T C T G G C C A G A G A G A G G T T T A T C T C T G T C T G C ----- 301

Db      121 A A C A A C C G G C A C C C A C A C C T T C T A C G T G G G A C T T A C A T T C T A T T G C T G T G G A C T 180
302 ----- A A A C C T C T C A C A C C T G C T T A G A G T G G G G C C T A T G C T T C A T C G G G T G G G A C 357

Qy      181 G T G A T G A T G T T T G A G G C T C T G G G G T G C T A T G G G G C A T C C A G A G A T C C C A G T T G C 240
Db      358 G T C A T A T G C T A T G G C C T T C C T G G G C T G A T G G G C C C G T C A C A G A G A G T C C G C T G 417

Qy      241 C T G G G A C G T T C T T C A C C T G C C T G T G A T C C T G T T T G C C T G T G A G A G T G G C T G A G G A T C 300
Db      418 C T G G G G C G T A C T T T G C T T C T G C T C T A T C T C A T T G C C C A G A G A G A G G C C G G G C C 477

Qy      301 T G G G G C T T C G T A A C A A A G A C A G A T G C C A G A G A T G T 339
Db      478 C T C T T C T A C T T C A A C A T G G G C A A G C A G C A C A G A G A T G 516

```

RESULT 15
US-08-705-771-8
Sequence 8, Application US/08705771
Patent No. 6034289
GENERAL INFORMATION:
APPLICANT: Paul Moore, Retner Gentz, Hongjin Ji,
APPLICANT: Jian Ni and Jing-Shan Hu
TITLE OF INVENTION: Human Genes, Sequences and
TITLE OF INVENTION: Expression Products
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CARILLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

ucleic - nucleic search, using sw model

December 9, 2003, 07:45:18 / Search time 1897 Seconds

(without alignments)
1194,884 Million cell updates/sec

US-10-035-914-1

1 aactaccgtctctctctc.....gaacagctcgtctactgag 682

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Number of hits satisfying chosen parameters: 4403344

num DB seq length: 0

num DB seq length: 2000000000

Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query	Match Length	ID	Description
1	534.8	78.4	1470 15 US-10-006-430-12
2	534.8	78.4	1470 15 US-10-006-430-12
3	534.8	78.4	1470 15 US-10-006-430-12
4	534.8	78.4	1470 15 US-10-006-430-12
5	501.6	73.5	1403 9 US-09-925-299-21
6	501.6	73.5	1403 9 US-09-925-299-21
7	499.8	73.3	1403 9 US-09-925-299-21
8	499.8	73.3	1403 9 US-09-925-299-21
9	499.8	73.3	1403 9 US-09-925-299-21
10	217.8	31.9	316 9 US-10-017-564-1349
11	217.8	31.9	316 9 US-10-017-564-1349
12	213.4	31.3	311 10 US-09-736-457-367
13	213.4	31.3	311 10 US-09-736-457-367
14	213.4	31.3	311 10 US-09-736-457-367
15	213.4	31.3	311 10 US-09-736-457-367
16	213.4	31.3	311 10 US-09-736-457-367
17	213.4	31.3	311 10 US-09-736-457-367
18	213.4	31.3	311 10 US-09-736-457-367
19	213.4	31.3	311 10 US-09-736-457-367
20	213.4	31.3	311 10 US-09-736-457-367
21	213.4	31.3	311 10 US-09-736-457-367
22	213.4	31.3	311 10 US-09-736-457-367
23	213.4	31.3	311 10 US-09-736-457-367
24	213.4	31.3	311 10 US-09-736-457-367
25	213.4	31.3	311 10 US-09-736-457-367
26	213.4	31.3	311 10 US-09-736-457-367
27	213.4	31.3	311 10 US-09-736-457-367
28	213.4	31.3	311 10 US-09-736-457-367
29	213.4	31.3	311 10 US-09-736-457-367
30	213.4	31.3	311 10 US-09-736-457-367
31	213.4	31.3	311 10 US-09-736-457-367
32	213.4	31.3	311 10 US-09-736-457-367
33	213.4	31.3	311 10 US-09-736-457-367
34	213.4	31.3	311 10 US-09-736-457-367
35	213.4	31.3	311 10 US-09-736-457-367
36	213.4	31.3	311 10 US-09-736-457-367
37	213.4	31.3	311 10 US-09-736-457-367
38	213.4	31.3	311 10 US-09-736-457-367
39	213.4	31.3	311 10 US-09-736-457-367
40	213.4	31.3	311 10 US-09-736-457-367
41	213.4	31.3	311 10 US-09-736-457-367
42	213.4	31.3	311 10 US-09-736-457-367
43	213.4	31.3	311 10 US-09-736-457-367
44	213.4	31.3	311 10 US-09-736-457-367
45	213.4	31.3	311 10 US-09-736-457-367

Query	Match Length	ID	Description
1	534.8	78.4	1470 15 US-10-006-430-12
2	534.8	78.4	1470 15 US-10-006-430-12
3	534.8	78.4	1470 15 US-10-006-430-12
4	534.8	78.4	1470 15 US-10-006-430-12
5	501.6	73.5	1403 9 US-09-925-299-21
6	501.6	73.5	1403 9 US-09-925-299-21
7	499.8	73.3	1403 9 US-09-925-299-21
8	499.8	73.3	1403 9 US-09-925-299-21
9	499.8	73.3	1403 9 US-09-925-299-21
10	217.8	31.9	316 9 US-10-017-564-1349
11	217.8	31.9	316 9 US-10-017-564-1349
12	213.4	31.3	311 10 US-09-736-457-367
13	213.4	31.3	311 10 US-09-736-457-367
14	213.4	31.3	311 10 US-09-736-457-367
15	213.4	31.3	311 10 US-09-736-457-367
16	213.4	31.3	311 10 US-09-736-457-367
17	213.4	31.3	311 10 US-09-736-457-367
18	213.4	31.3	311 10 US-09-736-457-367
19	213.4	31.3	311 10 US-09-736-457-367
20	213.4	31.3	311 10 US-09-736-457-367
21	213.4	31.3	311 10 US-09-736-457-367
22	213.4	31.3	311 10 US-09-736-457-367
23	213.4	31.3	311 10 US-09-736-457-367
24	213.4	31.3	311 10 US-09-736-457-367
25	213.4	31.3	311 10 US-09-736-457-367
26	213.4	31.3	311 10 US-09-736-457-367
27	213.4	31.3	311 10 US-09-736-457-367
28	213.4	31.3	311 10 US-09-736-457-367
29	213.4	31.3	311 10 US-09-736-457-367
30	213.4	31.3	311 10 US-09-736-457-367
31	213.4	31.3	311 10 US-09-736-457-367
32	213.4	31.3	311 10 US-09-736-457-367
33	213.4	31.3	311 10 US-09-736-457-367
34	213.4	31.3	311 10 US-09-736-457-367
35	213.4	31.3	311 10 US-09-736-457-367
36	213.4	31.3	311 10 US-09-736-457-367
37	213.4	31.3	311 10 US-09-736-457-367
38	213.4	31.3	311 10 US-09-736-457-367
39	213.4	31.3	311 10 US-09-736-457-367
40	213.4	31.3	311 10 US-09-736-457-367
41	213.4	31.3	311 10 US-09-736-457-367
42	213.4	31.3	311 10 US-09-736-457-367
43	213.4	31.3	311 10 US-09-736-457-367
44	213.4	31.3	311 10 US-09-736-457-367
45	213.4	31.3	311 10 US-09-736-457-367

ALIGNMENTS

Query	Match Length	ID	Description
1	534.8	78.4	1470 15 US-10-006-430-12
2	534.8	78.4	1470 15 US-10-006-430-12
3	534.8	78.4	1470 15 US-10-006-430-12
4	534.8	78.4	1470 15 US-10-006-430-12
5	501.6	73.5	1403 9 US-09-925-299-21
6	501.6	73.5	1403 9 US-09-925-299-21
7	499.8	73.3	1403 9 US-09-925-299-21
8	499.8	73.3	1403 9 US-09-925-299-21
9	499.8	73.3	1403 9 US-09-925-299-21
10	217.8	31.9	316 9 US-10-017-564-1349
11	217.8	31.9	316 9 US-10-017-564-1349
12	213.4	31.3	311 10 US-09-736-457-367
13	213.4	31.3	311 10 US-09-736-457-367
14	213.4	31.3	311 10 US-09-736-457-367
15	213.4	31.3	311 10 US-09-736-457-367
16	213.4	31.3	311 10 US-09-736-457-367
17	213.4	31.3	311 10 US-09-736-457-367
18	213.4	31.3	311 10 US-09-736-457-367
19	213.4	31.3	311 10 US-09-736-457-367
20	213.4	31.3	311 10 US-09-736-457-367
21	213.4	31.3	311 10 US-09-736-457-367
22	213.4	31.3	311 10 US-09-736-457-367
23	213.4	31.3	311 10 US-09-736-457-367
24	213.4	31.3	311 10 US-09-736-457-367
25	213.4	31.3	311 10 US-09-736-457-367
26	213.4	31.3	311 10 US-09-736-457-367
27	213.4	31.3	311 10 US-09-736-457-367
28	213.4	31.3	311 10 US-09-736-457-367
29	213.4	31.3	311 10 US-09-736-457-367
30	213.4	31.3	311 10 US-09-736-457-367
31	213.4	31.3	311 10 US-09-736-457-367
32	213.4	31.3	311 10 US-09-736-457-367
33	213.4	31.3	311 10 US-09-736-457-367
34	213.4	31.3	311 10 US-09-736-457-367
35	213.4	31.3	311 10 US-09-736-457-367
36	213.4	31.3	311 10 US-09-736-457-367
37	213.4	31.3	311 10 US-09-736-457-367
38	213.4	31.3	311 10 US-09-736-457-367
39	213.4	31.3	311 10 US-09-736-457-367
40	213.4	31.3	311 10 US-09-736-457-367
41	213.4	31.3	311 10 US-09-736-457-367
42	213.4	31.3	311 10 US-09-736-457-367
43	213.4	31.3	311 10 US-09-736-457-367
44	213.4	31.3	311 10 US-09-736-457-367
45	213.4	31.3	311 10 US-09-736-457-367

1 AAATACTGCTCTTCGTCCTTTCGATTTCGCTCTTCGAGCTGGAGAGGCGATCTAAGT 60
269 AACTACTCTCTCTGCTCTTCAATTGCTCTTCGCTGCTGAGAGGCTGATCTCGGCT 328
61 GTAGCTCTCTGCTGCTGATATATCAAGACACACAGCTCTGTACTTGGAACTGGGA 120
329 GTGGCCCTCTGCTCTCGCCATATACCGAGACCAACCACTCTCTGATCTTGGACTCTGGA 388

Query Match	78.4%	Score 534.8	DB 15	Length 1496
Best Local Similarity	86.5%	Prod. No. 1.3e-165		
Matches 599	Conservative 0	Mismatches 92	Indels 0	Gaps 0
QY	1	AAATACCTGCTCTTCGTCCTCAATTCCTTCGTGAGCTGAGGCGTGATCTTAACT	60	
	269	AAATACCTGCTCTTCGTCCTCAATTCCTTCGTGAGCTGAGGCGTGATCTTGGGT	328	

61 GTAGCTCTGTGGTGGTCAATGATCCAGACCAAGCCTGTGACTGAACTGGGA 120
 329 GTGGCCCTGTGGCTCGCCATGACCCGAGACCACTCTCTGTATGTGAGCTGGGA 388
 121 AACAAACCGGACCCCAACACCTTTAGTGGGATCTCAATTCATTCGTGGAGCT 180
 389 GACAAAGCCGCGCCCAACCTTTATGTAGGATCTCAATTCCTCATTCCTGTGGGCGCT 448
 181 GTGATGATGTTTGTAGGCTTCGAGGGGTCTATGAGGGCCATCCAGAGTCCAGTCTG 240
 449 GTCAATGATGTTGTGGCTTCCTGAGCTCTAGAGGGCCATCCAGAGTCCAGTCTG 508
 241 CTGGGAGCGTTCTTCACTGCTGTGATCTGTTGCTGTGAGGCTGAGGCTG 300
 509 CTGGGAGCGTTCTTCACTGCTGTGATCTGTTGCTGTGAGGCTGAGGCTG 568
 301 TGGGGCTTGTAAACAAAGACAGATGCGCAAGATGTGAGAGCGTTCTATGACCAAGCC 360
 569 TGGGGCTTGTAAACAAAGACAGATGCGCAAGATGTGAGAGCGTTCTATGACCAAGCC 628
 361 CTTCAGCAAGCTGTATGATGATGATGCAATGCAAGGCTGTGTGAAAGCTTTC 420
 629 CTACAGCAGCGCTGTGTGATGATGATGCAAGGCTGTGTGAAAGCTTTC 688
 421 CATGAGAGCTGCACTGTTGTGCTCCAGACGCACTGACCACTGACCACTGAC 480
 689 CACGAGAGCTGCACTGTTGTGCTCCAGACGCACTGACCACTGACCACTGAC 748
 481 AGGAAACCGCTGTGCTGCTCCAGACGCACTGACCACTGACCACTGACCACTG 540
 749 AAGAAACCGCTGTGCTGCTCCAGACGCACTGACCACTGACCACTGACCACTG 808
 541 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 809 CACGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 868
 601 GTGTAGCTGTCAATTATGATCTTTGAGATGATCTGAGATGATGATGATGATGATG 660
 869 GTGTAGCTGTCAATTATGATCTTTGAGATGATCTGAGATGATGATGATGATGATG 928
 661 CGGAAACGCTCGGTGATGAG 682
 929 CGGAAACGCTCGGTGATGAG 950

1-006-430-3
 Application US/10006430
 Application No. US20030113914A1

INVENTOR: Kenneth Dobie
 APPLICANT: Mark J. Graham

TITLE OF INVENTION: ANTISENSE MODULATION OF CD81 EXPRESSION

REFERENCE: RTS-0341
 PRESENT APPLICATION NUMBER: US/10/006,430
 FILING DATE: 2001-12-10

MEMBER OF SEQ ID NOS: 90
 ID NO 3

LENGTH: 1496
 TYPE: DNA

ORGANISM: Homo sapiens
 NAME/KEY: CDS
 LOCATION: (239) ... (949)
 -006-430-3

1 AAATACCTGCTCTTCTTCTTCAATTTCTTCTGCTGCTGAGGCTGATCTTAACT 60
 269 AAGTACCTGCTCTTCTTCTTCAATTTCTTCTGCTGCTGAGGCTGATCTTAACT 328

QY 61 GTAGCTCTGTGGTGGTCAATGATCCAGACCAAGCCTGTGACTGAACTGGGA 120
 Db 329 GTGGCCCTGTGGCTCGCCATGACCCGAGACCACTCTCTGTATGTGAGCTGGGA 388
 QY 121 AACAAACCGGACCCCAACACCTTTAGTGGGATCTCAATTCATTCGTGGAGCT 180
 Db 389 GACAAAGCCGCGCCCAACCTTTATGTAGGATCTCAATTCCTCATTCCTGTGGGCGCT 448
 QY 181 GTGATGATGTTTGTAGGCTTCGAGGGGTCTATGAGGGCCATCCAGAGTCCAGTCTG 240
 Db 449 GTCAATGATGTTGTGGCTTCCTGAGCTCTAGAGGGCCATCCAGAGTCCAGTCTG 508
 QY 241 CTGGGAGCGTTCTTCACTGCTGTGATCTGTTGCTGTGAGGCTGAGGCTG 300
 Db 509 CTGGGAGCGTTCTTCACTGCTGTGATCTGTTGCTGTGAGGCTGAGGCTG 568
 QY 301 TGGGGCTTGTAAACAAAGACAGATGCGCAAGATGTGAGAGCGTTCTATGACCAAGCC 360
 Db 569 TGGGGCTTGTAAACAAAGACAGATGCGCAAGATGTGAGAGCGTTCTATGACCAAGCC 628
 QY 361 CTTCAGCAAGCTGTATGATGATGATGCAATGCAAGGCTGTGTGAAAGCTTTC 420
 Db 629 CTACAGCAGCGCTGTGTGATGATGATGCAAGGCTGTGTGAAAGCTTTC 688
 QY 421 CATGAGAGCTGCACTGTTGTGCTCCAGACGCACTGACCACTGACCACTGAC 480
 Db 689 CACGAGAGCTGCACTGTTGTGCTCCAGACGCACTGACCACTGACCACTGAC 748
 QY 481 AGGAAACCGCTGTGCTGCTCCAGACGCACTGACCACTGACCACTGACCACTG 540
 Db 749 AAGAAACCGCTGTGCTGCTCCAGACGCACTGACCACTGACCACTGACCACTG 808
 QY 541 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 Db 809 CACGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 868
 QY 601 GTGTAGCTGTCAATTATGATCTTTGAGATGATCTGAGATGATGATGATGATGATG 660
 Db 869 GTGTAGCTGTCAATTATGATCTTTGAGATGATCTGAGATGATGATGATGATGATG 928
 QY 661 CGGAAACGCTCGGTGATGAG 682
 Db 929 CGGAAACGCTCGGTGATGAG 950

RESULT 5
 US-09-925-299-21

Sequence 21, Application US/09925299
 Patent No. US20020055627A1

GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FIL REFERENCE: PA102
 CURRENT APPLICATION NUMBER: US/09/925,299

PRIOR FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05883
 PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270
 NUMBER OF SEQ ID NOS: 1556
 SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 21
 LENGTH: 1403
 TYPE: DNA

ORGANISM: Homo sapiens
 NAME/KEY: misc feature
 LOCATION: (1370)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: misc feature
 LOCATION: (1386)
 OTHER INFORMATION: n equals a,t,g, or c

ent No. US20020076715A1
 3RAL INFORMATION:
 PLICANT: Benson, Darin R.
 PLICANT: Lodes, Michael J.
 PLICANT: Mitcham, Jennifer L.
 PLICANT: King, Gordon B.
 FILE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
 FILE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
 SE REFERENCE: 210121.466C3
 PRESENT APPLICATION NUMBER: US/09/876,889
 PRESENT FILING DATE: 2001-06-06
 NUMBER OF SEQ ID NOS: 353
 TMAR: FastSeq for Windows Version 3.0
 ID NO 303
 LENGTH: 436
 TYPE: DNA
 ORGANISM: Homo sapien
 -876-889-303

Y Match 39.6%; Score 270; DB 9; Length 436;
 : Local Similarity 100.0%; Pred. No. 1.9e-78;
 :hes 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

413 AGACTTTCATGAGCTCACTGTTGGCTCCAGCACTGACCACTGACTACCA 472
 |||||
 1 AGACTTTCATGAGCTCACTGTTGGCTCCAGCACTGACCACTGACTACCA 60
 |||||
 473 CCATACGAGCAACAGCTGTGCTCCAGGCGCAACATACACCCCTTACTGACG 532
 |||||
 61 CCATACGAGCAACAGCTGTGCTCCAGGCGCAACATACACCCCTTACTGACG 120
 |||||
 533 AGATTGTATCAAGAAATGATGAGCTCTTCTTGGGAAGCTGTACTGAAATG 592
 |||||
 121 AGATTGTATCAAGAAATGATGAGCTCTTCTTGGGAAGCTGTACTGAAATG 180
 |||||
 593 CAGCATGTGGAGCTGTATATGATCTTGAAGTATTCGAGCAAGTGTGTGT 652
 |||||
 181 CAGCATGTGGAGCTGTATATGATCTTGAAGTATTCGAGCAAGTGTGTGT 240
 |||||
 653 GTGGCATCCGGAACAGCTCCGTACTGAG 682
 |||||
 241 GTGGCATCCGGAACAGCTCCGTACTGAG 270
 |||||

: 10
 : -777-564-1349/c
 :ence 1349, Application US/09777564
 :ent No. US20020022591A1
 :RAL INFORMATION:
 PLICANT: Algate, Paul A.
 PLICANT: Mamion, Jane
 FILE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 FILE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 SE REFERENCE: 210121.493
 PRESENT APPLICATION NUMBER: US/09/777,564
 PRESENT FILING DATE: 2001-02-05
 NUMBER OF SEQ ID NOS: 1730
 TMAR: FastSeq for Windows Version 4.0
 ID NO 1349
 LENGTH: 316
 TYPE: DNA
 ORGANISM: Homo sapiens
 -777-564-1349

Y Match 31.9%; Score 217.8; DB 9; Length 316;
 : Local Similarity 82.7%; Pred. No. 2.7e-61;
 :hes 249; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

382 GATGATCCCAATATCCCAAGCTGTGTGGAAGCTTTCATGAGAGCTCACTGTGT 441
 |||||
 316 GATGATCCCAATATCCCAAGCTGTGTGGAAGCTTTCATGAGAGCTCACTGTGT 257
 |||||
 442 GGCTCCAGCACTGACCACTGATCTTACCAACATATGAGGAACAGCTGTGTCTCA 501
 |||||
 256 GGCTCCAGCACTGACCACTGATCTTACCAACATATGAGGAACAGCTGTGTCTCA 197
 |||||
 502 GGCGCAACATATCAACCCCTTACTGAGCAAGATTTGTCATGAGAAATGATGAGCTC 561
 |||||
 136 GGCGCAACATATCAACCCCTTACTGAGCAAGATTTGTCATGAGAAATGATGAGCTC 137
 |||||
 562 TTCTCTGGAAGCTGTACTCTCATTTGGAATTCAGCAATTCGATGTGATCATTAATGATC 621
 |||||
 136 TTCTCTGGAAGCTGTACTCTCATTTGGAATTCAGCAATTCGATGTGATCATTAATGATC 77
 |||||
 622 TTGAGATGATTCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681
 |||||
 76 TTGAGATGATTCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 17
 |||||
 682 G 682
 |
 16 G 16

442 GGCTCCAGCACTGACCACTGATCAACCATATGAGGAACAGCTGTGTCTCA 501
 |||||

Db 256 GGCTCCAGCACTGACCACTGATCAACCATATGAGGAACAGCTGTGTCTCA 197
 |||||
 Oy 502 GGCGCAACATATCAACCCCTTACTGAGCAAGATTTGTCATGAGAAATGATGAGCTC 561
 |||||
 Db 136 GGCGCAACATATCAACCCCTTACTGAGCAAGATTTGTCATGAGAAATGATGAGCTC 137
 |||||
 Oy 562 TTCTCTGGAAGCTGTACTCTCATTTGGAATTCAGCAATTCGATGTGATCATTAATGATC 621
 |||||
 Db 136 TTCTCTGGAAGCTGTACTCTCATTTGGAATTCAGCAATTCGATGTGATCATTAATGATC 77
 |||||
 Oy 622 TTGAGATGATTCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681
 |||||
 Db 76 TTGAGATGATTCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 17
 |||||
 Oy 682 G 682
 |
 Db 16 G 16

RESULT 11
 US-10-015-219-1349/c
 : Sequence 1349, Application US/10015219
 : Publication No. US20030008299A1
 : GENERAL INFORMATION:
 : APPLICANT: Algate, Paul A.
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 : FILE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 : SE REFERENCE: 210121.493C1
 : CURRENT APPLICATION NUMBER: US/10/015.219
 : PRESENT FILING DATE: 2002-03-02
 : NUMBER OF SEQ ID NOS: 1739
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1349
 : LENGTH: 316
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-10-015-219-1349

Query Match 31.9%; Score 217.8; DB 15; Length 316;
 : Local Similarity 82.7%; Pred. No. 2.7e-61;
 : Matches 249; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Oy 382 GATGATCCCAATATCCCAAGCTGTGTGGAAGCTTTCATGAGAGCTCACTGTGT 441
 |||||
 Db 316 GATGATCCCAATATCCCAAGCTGTGTGGAAGCTTTCATGAGAGCTCACTGTGT 257
 |||||
 Oy 442 GGCTCCAGCACTGACCACTGATCTTACCAACATATGAGGAACAGCTGTGTCTCA 501
 |||||
 Db 256 GGCTCCAGCACTGACCACTGATCTTACCAACATATGAGGAACAGCTGTGTCTCA 197
 |||||
 Oy 502 GGCGCAACATATCAACCCCTTACTGAGCAAGATTTGTCATGAGAAATGATGAGCTC 561
 |||||
 Db 136 GGCGCAACATATCAACCCCTTACTGAGCAAGATTTGTCATGAGAAATGATGAGCTC 137
 |||||
 Oy 562 TTCTCTGGAAGCTGTACTCTCATTTGGAATTCAGCAATTCGATGTGATCATTAATGATC 621
 |||||
 Db 136 TTCTCTGGAAGCTGTACTCTCATTTGGAATTCAGCAATTCGATGTGATCATTAATGATC 77
 |||||
 Oy 622 TTGAGATGATTCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681
 |||||
 Db 76 TTGAGATGATTCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 17
 |||||
 Oy 682 G 682
 |
 Db 16 G 16

RESULT 12
 US-09-736-457-367/c
 : Sequence 367, Application US/09736457
 : Patent No. US20020168637A1
 : GENERAL INFORMATION:

APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaletanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Pan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
Q ID NO 367
LENGTH: 311
TYPE: DNA
ORGANISM: Homo sapien
9-736-457-367
Query Match 31.3%; Score 213.4; DB 10; Length 311;
Best Local Similarity 82.7%; Pred. No. 7.6e-60;
Matches 244; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
388 GCCACATATCCAGAGCTGTGTGTAAGACTTTCATGAGACGCTCAACTGTGTGCTCC 447
310 GCCAACAAAGCCAGAGCTGTGTGTAAGACTTTCATGAGACGCTCAACTGTGTGCTCC 251
448 AACGACTGACCACTGACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 507
250 AGCACTGACCACTGACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 191
508 AACATCTACACCCCTTACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 567
190 AACATCTACACCCCTTACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 131
568 GGGAGCTGTACTCTTATGAAATTCAGAGCATTGTGAGCTGTATATGATCTTTGAG 627
130 GGGAGCTGTACTCTTATGAAATTCAGAGCATTGTGAGCTGTATATGATCTTTGAG 71
628 ATGATCTGAGCATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 682
70 ATGATCTGAGCATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 16
UT 13
9-902-941-367/c
Sequence 367; Application US/09902941
Ent No. US20020172952A1
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Henderison, Robert A.
APPLICANT: Wang, Tonglong
APPLICANT: Mafanabe, Yoshihiro
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Retter, Marc W.
APPLICANT: Marnerakis, Margarita
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Bangur, Chaletanya S.
APPLICANT: Mcnabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C17
CURRENT APPLICATION NUMBER: US/09/902,941
CURRENT FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 2002
SOFTWARE: FastSeq for Windows Version 4.0
Q ID NO 367
LENGTH: 311

TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-941-367
Query Match 31.3%; Score 213.4; DB 10; Length 311;
Best Local Similarity 82.7%; Pred. No. 7.6e-60;
Matches 244; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
388 GCCAACAAAGCCAGAGCTGTGTGTAAGACTTTCATGAGACGCTCAACTGTGTGCTCC 447
310 GCCAACAAAGCCAGAGCTGTGTGTAAGACTTTCATGAGACGCTCAACTGTGTGCTCC 251
448 AACGACTGACCACTGACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 507
250 AGCACTGACCACTGACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 191
508 AACATCTACACCCCTTACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 567
190 AACATCTACACCCCTTACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 131
568 GGGAGCTGTACTCTTATGAAATTCAGAGCATTGTGAGCTGTATATGATCTTTGAG 627
130 GGGAGCTGTACTCTTATGAAATTCAGAGCATTGTGAGCTGTATATGATCTTTGAG 71
628 ATGATCTGAGCATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 682
70 ATGATCTGAGCATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 16
RESULT 14
US-09-849-626-367/c
Sequence 367; Application US/09849626
Publication No. US20020197669A1
GENERAL INFORMATION:
APPLICANT: Bangur, Chaletanya
APPLICANT: Fanger, Gary
APPLICANT: Wang, Aijun
APPLICANT: Wang, Tonglong
APPLICANT: Switzer, Anne
APPLICANT: McNeill, Patricia
APPLICANT: Clapper, Jonathan
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C16
CURRENT APPLICATION NUMBER: US/09/849,626
CURRENT FILING DATE: 2001-05-03
NUMBER OF SEQ ID NOS: 1926
SOFTWARE: FastSeq for Windows Version 3.0
Q ID NO 367
LENGTH: 311
TYPE: DNA
ORGANISM: Homo sapien
US-09-849-626-367
Query Match 31.3%; Score 213.4; DB 10; Length 311;
Best Local Similarity 82.7%; Pred. No. 7.6e-60;
Matches 244; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
388 GCCAACAAAGCCAGAGCTGTGTGTAAGACTTTCATGAGACGCTCAACTGTGTGCTCC 447
310 GCCAACAAAGCCAGAGCTGTGTGTAAGACTTTCATGAGACGCTCAACTGTGTGCTCC 251
448 AACGACTGACCACTGACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 507
250 AGCACTGACCACTGACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 191
508 AACATCTACACCCCTTACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 567
190 AACATCTACACCCCTTACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 131
568 GGGAGCTGTACTCTTATGAAATTCAGAGCATTGTGAGCTGTATATGATCTTTGAG 627
130 GGGAGCTGTACTCTTATGAAATTCAGAGCATTGTGAGCTGTATATGATCTTTGAG 71

628 ATGATTCGACATGCTGCTGCTGCGATCCGGAACAGCTCCGTGACTGAG 682
70 ATGATTCGACATGCTGCTGCTGCGATCCGGAACAGCTCCGTGACTGAG 16

T 15
-476-300-367/c
uence 367, Application US/09476300
lication No. US20030125245A1
RAL INFORMATION:
PLICANT: Bang, Tonglong
TLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
LE REFERENCE: 210121.478C3
RRENT APPLICATION NUMBER: US/09/476,300
RRENT FILING DATE: 1999-12-30
MBER OF SEQ ID NOS: 785
FTWARE: FastSeq for Windows Version 3.0
ID NO 367
ENGTH: 311
YPE: DNA
GANISM: Homo sapien
-476-300-367

31.3%; Score 213.4; DB 11; Length 311;
c Local Similarity 82.7%; Pred. No. 7.6e-60;
ches 244; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
388 GCCAACAATGCCAAGCTGTGTGTAAGACTTTCATGAGAGCTCAACTGTGTGCTCC 447
310 GCCAACAACGCCAAGCTGTGTGTAAGACTTTCATGAGAGCTCAACTGTGTGCTCC 251
448 AACGCACTGACCACTGACTACCATCTGAGAGAGAGAGCTGTGTGCTGAGAGGCGC 507
250 AGCACTGACCTGCTTGAACCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 191
508 AACATCTCACTCCCTTACTGACAGATGTCATGAGAAATGATGAGCTCTTCTCT 567
190 AACATCTCACTCACTCTTCAAGAGAGAGCTGCAACAGAGATGATGAGCTCTTCTCC 131
568 GGGAAAGCTGACTCATGGAATTGACGCAATGTTGTTAGCTGTCAATGATCTTTGAG 627
130 GGGAAAGCTGACTCATGGAATTGACGCAATGTTGTTAGCTGTCAATGATCTTTGAG 71
628 ATGATTCGACATGCTGCTGCTGCGATCCGGAACAGCTCCGTGACTGAG 682
70 ATGATTCGACATGCTGCTGCTGCGATCCGGAACAGCTCCGTGACTGAG 16

1 completed: December 9, 2003, 10:28:13
Time: 1901 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

nucleic - nucleic search, using sw model

on: December 9, 2003, 01:37:00 ; Search time 2618 Seconds
(without alignments)
6331.418 Million cell updates/sec

US-10-035-914-1

ect score: 682
ence: 1 aaataccgctcctcctc.....gaacagctcctgctacgag 682

ing table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

ched: 22781392 seqs, 12152238056 residues

1 number of hits satisfying chosen parameters: 45562784

num DB seq length: 0
num DB seq length: 200000000

-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

base :

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estcov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vit:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vit:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

lc	Query	Score	Match	Length	DB	ID	Description
1	682	100.0	885	12	BI695471	BI695471	603346477
2	675.6	99.1	1169	13	BUS15266	AGENCOURT	BUS15266
3	675.4	99.0	918	13	BQ900007	AGENCOURT	BQ900007
4	661	96.9	881	14	CB196088	AGENCOURT	CB196088

5	660	96.8	698	12	BI157227	BI157227	602923435
6	657.4	96.4	777	12	BI653596	BI653596	603300230
7	655.6	96.1	956	13	BQ714590	AGENCOURT	BQ714590
8	650.8	95.4	933	13	BUS23591	AGENCOURT	BUS23591
9	645	94.6	852	14	CB556103	AGENCOURT	CB556103
10	641.2	94.0	759	12	BI113178	AGENCOURT	BI113178
11	640	93.8	759	12	BI695388	AGENCOURT	BI695388
12	639.4	93.8	703	13	BQ886993	AGENCOURT	BQ886993
13	639	93.7	642	10	BF531191	AGENCOURT	BF531191
14	633.4	92.9	846	12	BI694804	AGENCOURT	BI694804
15	627.2	92.0	806	10	BF181041	AGENCOURT	BF181041
16	619	90.8	958	13	BQ934764	AGENCOURT	BQ934764
17	613.8	90.0	1002	13	BUS59889	AGENCOURT	BUS59889
18	612.4	89.8	647	10	BF138851	AGENCOURT	BF138851
19	611.4	89.6	945	13	BU147770	AGENCOURT	BU147770
20	607.2	89.0	918	10	BF139437	AGENCOURT	BF139437
21	605.8	88.8	884	13	BQ886347	AGENCOURT	BQ886347
22	603	88.4	867	13	BU505471	AGENCOURT	BU505471
23	601	88.1	1030	10	BF783079	AGENCOURT	BF783079
24	597	87.5	988	12	BI453448	AGENCOURT	BI453448
25	596	87.4	737	12	BI693368	AGENCOURT	BI693368
26	593	87.0	872	13	BUS17509	AGENCOURT	BUS17509
27	592.2	86.8	851	12	BI690870	AGENCOURT	BI690870
28	586.2	86.0	818	12	BI694417	AGENCOURT	BI694417
29	584	85.6	937	13	BQ937206	AGENCOURT	BQ937206
30	574.6	84.3	847	12	BI110858	AGENCOURT	BI110858
31	571.8	83.8	866	12	BI689533	AGENCOURT	BI689533
32	570	83.6	931	13	BQ947699	AGENCOURT	BQ947699
33	569	83.4	996	13	BQ886352	AGENCOURT	BQ886352
34	563.2	82.6	903	12	BI079904	AGENCOURT	BI079904
35	562.4	82.5	744	12	BI854539	AGENCOURT	BI854539
36	558.6	81.9	735	10	BG668442	DRABYCO7	BG668442
37	557	81.7	921	13	BQ714029	AGENCOURT	BQ714029
38	555.8	81.5	892	13	BQ219107	AGENCOURT	BQ219107
39	553.2	81.1	844	12	BI146915	AGENCOURT	BI146915
40	552.2	81.0	925	13	BQ926886	AGENCOURT	BQ926886
41	552	80.9	917	10	BF15104	AGENCOURT	BF15104
42	552	80.9	956	13	BUS18267	AGENCOURT	BUS18267
43	551.4	80.9	933	12	BI695295	AGENCOURT	BI695295
44	546.6	80.1	824	13	BUS58495	AGENCOURT	BUS58495
45	536.4	78.7	991	13	BUS39845	AGENCOURT	BUS39845

ALIGNMENTS

RESULT 1
LOCUS BI695471 885 bp mRNA linear EST 18-SEP-2001
DEFINITION 603346477F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5374171 5',
mRNA sequence.
ACCESSION BI695471
VERSION BI695471.1 GI:15658100
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 885)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLNL1952 row: 1 column: 20
High quality sequence stop: 816.

421 CATGAGCGCTCACTGTTGAGCTCCAAAGCACTGACCACTGATCAACCAATATCTG 488
679 CATGAGACGCTCACTGTTGAGCTCCAAAGCACTGACCACTGATCAACCAATATCTG 738
481 AGGAACAGCTCTGTCTCCCTCAGGCGGCAACATCTCAACCCCTTACTGCAAGCAAGATTGT 540
739 AGGAACAGCTCTGTCTCCCTCAGGCGGCAACATCTCAACCCCTTACTGCAAGCAAGATTGT 798
541 CATCAGAAAAATCGATGAGCTCTTCTCTCGGAAGCTGACCTCATTTGAAATTGCAAGCAATT 600
799 CATCAGAAAAATCGATGAGCTCTTCTCTCGGAAGCTGACCTCATTTGAAATTGCAAGCAATT 858
601 GTGTAGCTGTCAATTATGATCTTTGAGATGATTTCTGAGCAATGCTGTGCTGTGCAATC 660
859 GTGTAGCTGTCAATTATGATCTTTGAGATGATTTCTGAGCAATGCTGTGCTGTGCAATC 918
661 CGGAACAGCTCCGTTACTGAG 682
919 CGGAACAGCTCCGTTACTGAG 940

ILT 3
 10007
 BQ900007 918 bp mRNA linear EST 16-AUG-2002
 INITION AGENCOCURT 8747038 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6532966
 5' , mRNA sequence.
 BQ900007
 BQ900007.1 GI:22292033
 ORDS EST.

CR. Mus musculus (house mouse)
 GENISM. Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scturognathi; Muridae; Murinae; Mus
 1 (bases 1 to 918)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.

CDNA library preparation: Keesden, InVivoGen Corp
CDNA library arrayed by: The I.M.A.G.E. Consortium
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13883 row: k column: 16
High quality sequence start: 577.

1. .918

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6392967"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH_MGC_130"
/name="Organ: oocytes; Vector: PCMV-SPOrt6.1.ccd;
Site 1: Score; Site 2: NotI; Cloned undirectionally.
Size: 0.190 kb. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is a
NH_MGC Library."
COUNT
194 a 257 c 245 g 220 t 2 others

```

```

1  AATAACCGCTCTGCTTCATATTTGCTCTTCTGCTGCTGCTGAGCGCTGATCTAGGT 60
150 AATAACCGCTCTGCTTCATATTTGCTCTTCTGCTGCTGCTGAGCGCTGATCTAGGT 209

```

QY	61	GTAGCTCTGTGGTTCGTCATCATCAAGACACACAGCTGTGTATACCTGTGAACCTGGAA	120
Db	210	GTAGCTCTGTGGTTCGTCATCATCAAGACACACAGCTGTGTATACCTGTGAACCTGGAA	269
QY	121	MACAAAACGGGCAACCCACACCTTCTACGTGGGCACTCAATCTCATTTGCTGTGGAGCT	180
Db	270	MACAAAACGGGCAACCCACACCTTCTACGTGGGCACTCAATCTCATTTGCTGTGGAGCT	329
QY	181	GTGATGATGTTTGTAGGGCTTCCTGGGGGGCTATATGGGGGCATTCAGAGATCCAGTGTCTTG	240
Db	330	GTGATGATGTTTGTAGGGCTTCCTGGGGGGCTATATGGGGGCATTCAGAGATCCAGTGTCTTG	389
QY	241	CTGGGGAAGTTCTTCAACCTGCCTTGTGATTCCTGTTTGCTGTGAGGTGGCTCAGGATTC	300
Db	390	CTGGGGAAGTTCTTCAACCTGCCTTGTGATTCCTGTTTGCTGTGAGGTGGCTCAGGATTC	449
QY	301	TGGGGCTTCGTAAACAAAGACACATTCGCAAGATGTGAGACAGTTCTATGACCAAGCC	360
Db	450	TGGGGCTTCGTAAACAAAGACACATTCGCAAGATGTGAGACAGTTCTATGACCAAGCC	509
QY	361	CTTCAGCAGCTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	420
Db	510	CTTCAGCAGCTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	569
QY	421	CATGAGACGCTCAACTGTTGTGGGCTCCAAAGGCACTGACCACTGACTACCACTAATCTG	480
Db	570	CATGAGACGCTCAACTGTTGTGGGCTCCAAAGGCACTGACCACTGACTACCACTAATCTG	629
QY	481	AGGAACAGCCTGTGTCCCTCAAGGGGGGAAATACTCAACCCCTTACTGACGAAATGTT	540
Db	630	AGGAACAGCCTGTGTCCCTCAAGGGGGGAAATACTCAACCCCTTACTGACGAAATGTT	689
QY	541	CATCAGAAATCGATGAGCTCTTCTCTGGGAAAGCTGTACTCATTTGGAATTTGACGCAAT	600
Db	690	CATCAGAAATCGATGAGCTCTTCTCTGGGAAAGCTGTACTCATTTGGAATTTGACGCAAT	749
QY	601	GTGTTACTGTGCATTATGATCTTTTGAATGATTTCTGAGCATGTGTGTGTGTGGCATTC	660
Db	750	GTGTTACTGTGCATTATGATCTTTTGAATGATTTCTGAGCATGTGTGTGTGTGGCATTC	809
QY	661	CGGAACAGCTCCGTTGA 677	
Db	810	CGGAACAGCTCCGTTGA 826	

RESULT 4					
CB196088					
LOCUS					
DEFINITION	CB196088	881 bp	RNA	linear	EST 05-FEB-2003
ACCESSION	IMAGE:CCURT_11259825	NIH MGC_135	Mus musculus	cDNA clone	
VERSION	CB196088				
KEYWORDS	CB196088.1	GI:28223311			
SOURCE	EST.				
ORGANISM	Mus musculus	(house mouse)			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (bases 1 to 881)				
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/ .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished				
	Contact: Robert Strausberg, Ph.D.				

Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
Plate: NDA00039 row: 1 column: 08
High quality sequence start: 18
High quality sequence stop: 711.

RES Location/Qualifiers
Source 1. 881
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30137839"
/lab_host="DH10B (phage-resistant)"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5,
and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dT. cDNA
enrichment: >1k bp, Average insert size 1.6k bp.
Normalization (Cot value): 7.5 kb. Priming sequence:
5'-GACTGTTCTAGATGCGAGCGCGCC(7)3' Tissue contributed by
David Rowe. Library constructed by Reegen, Invitrogen
Corp."

COUNT 174 a 211 c 231 g 265 t
Match 96.9%; Score 661; DB 14; Length 881;
Local Similarity 100.0%; Pred. No. 1.6e-164; Indels 0; Gaps 0;
ches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

22 AATTTCGTTCTTGGCTGGCTGGAGGCTGGATCCTAGGTAGCTGTGTGGTGGTCTAT 81
27 AATTTCGTTCTTGGCTGGCTGGAGGCTGGATCCTAGGTAGCTGTGTGGTGGTCTAT 86
82 GATCCAGACCAACCAAGCTGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCT 141
87 GATCCAGACCAACCAAGCTGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCT 146
142 TTCTACGGGGGCACTCTCATCTGCTGTGGAGCTGTGATGATGATGATGATGATGATG 201
147 TTCTACGGGGGCACTCTCATCTGCTGTGGAGCTGTGATGATGATGATGATGATGATG 206
202 CTGGGGGCTGTAAGGGGCACTCAGAGGCTCAGAGGCTCAGAGGCTCAGAGGCTCAGAG 261
207 CTGGGGGCTGTAAGGGGCACTCAGAGGCTCAGAGGCTCAGAGGCTCAGAGGCTCAGAG 266
262 CTTGTGATCTGTTGGCTGTGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 321
267 CTTGTGATCTGTTGGCTGTGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 326
322 CAGATCGCAAGAGTGAAGCAGTCTATGATGATGATGATGATGATGATGATGATGATG 381
327 CAGATCGCAAGAGTGAAGCAGTCTATGATGATGATGATGATGATGATGATGATGATG 386
382 GATGATGCAACAATGCAAGGCTGTGGTGAAGCTTTCATGAGAGCTCAACTGTTGT 441
387 GATGATGCAACAATGCAAGGCTGTGGTGAAGCTTTCATGAGAGCTCAACTGTTGT 446
442 GGTTCGCAAGCACTGACCACTGACCACTGACCACTGACCACTGACCACTGACCACTG 501
447 GGTTCGCAAGCACTGACCACTGACCACTGACCACTGACCACTGACCACTGACCACTG 506
502 GGGGGCAACATCTCACTCCCTTACTGAGAGCAAGTGTGATGATGATGATGATGATGATG 561
507 GGGGGCAACATCTCACTCCCTTACTGAGAGCAAGTGTGATGATGATGATGATGATGATG 566
562 TTCTCTGGGAAGCTGTAAGCTGTAAGTGAATGACCACTGTAAGCTGTAAGCTGTAAG 621
567 TTCTCTGGGAAGCTGTAAGCTGTAAGTGAATGACCACTGTAAGCTGTAAGCTGTAAG 626
622 TTCTCTGGGAAGCTGTAAGCTGTAAGTGAATGACCACTGTAAGCTGTAAGCTGTAAG 681
627 TTCTCTGGGAAGCTGTAAGCTGTAAGTGAATGACCACTGTAAGCTGTAAGCTGTAAG 686
682 G 682
687 G 687

RESULT 5
BI157227 698 bp mRNA linear EST 05-JUL-2001
LOCUS 602923435F1_NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5063432 5',
DEFINITION mRNA sequence.
ACCESSION BI157227
VERSION BI157227
KEYWORDS EST.
SOURCE BI157227.1 GI:14617228
ORGANISM Mus musculus (house mouse)
REFERENCE 1. (bases 1 to 698)
NIH-MGC <http://mgc.nci.nih.gov/>.
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
Tissue Procurement: Lothar Hemmighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
CDNA Library Arayed by: The I.M.A.G.E. Consortium (LNL)
DNA Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.jnl.gov>
Plate: LLM1171 row: f column: 09
High quality sequence stop: 696.
Location/Qualifiers
1. 698
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5063432"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_11b="NCI_CGAP_Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hemmighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

FEATURES
Source
Location/Qualifiers
1. 698
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5063432"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_11b="NCI_CGAP_Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hemmighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

BASE COUNT 149 a 180 c 190 g 179 t
ORIGIN

Query Match 96.8%; Score 660; DB 12; Length 698;
Best Local Similarity 100.0%; Pred. No. 2.6e-164; Indels 0; Gaps 0;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

23 AATTTCGTTCTTGGCTGGCTGGAGGCTGGATCCTAGGTAGCTGTGTGGTGGTCTAT 82
1 AATTTCGTTCTTGGCTGGCTGGAGGCTGGATCCTAGGTAGCTGTGTGGTGGTCTAT 86
83 ATCCAGAGCAACCAAGCTGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCT 142
61 ATCCAGAGCAACCAAGCTGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCT 120
143 TCTACGGGGGCACTCTCATCTGCTGTGGAGCTGTGATGATGATGATGATGATGATG 202
121 TCTACGGGGGCACTCTCATCTGCTGTGGAGCTGTGATGATGATGATGATGATGATG 180
203 TGGGGTGTATGAGGCACTCAGAGGCTCCAGTGTGCTGAGGAGCGTTCTTCACTGACC 262
181 TGGGGTGTATGAGGCACTCAGAGGCTCCAGTGTGCTGAGGAGCGTTCTTCACTGACC 240
263 TTGTGATCTGTTGGCTGTGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 322
241 TTGTGATCTGTTGGCTGTGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 300
323 AGATGCGCAAGAGTGAAGCAGTCTATGATGATGATGATGATGATGATGATGATGATG 382

|||||
301 AGATCCCAAGATGTAAGCAGTTCTATACAGGCCCTTCAGCAAGCTGTATGATG 360
383 ATGATGCAACATGTCAGAGGCTGTGTGAAGACTTTCATGAGAGCGCTCACTGTG 442
361 ATGATCCCAAGATGTAAGCAGTTCTATACAGGCCCTTCAGCAAGCTGTATG 420
443 GCTCAACGCACTGACGACATGACTACCACTACATGAGGAAACAGCTGTGCTCAG 502
421 GCTCAACGCACTGACGACATGACTACCACTACATGAGGAAACAGCTGTGCTCAG 480
503 GCGGCAACATCTACCCCTTATCTGACAGAGATGTGATCAGAAATGATGAGCTCT 562
481 GCGGCAACATCTACCCCTTATCTGACAGAGATGTGATCAGAAATGATGAGCTCT 540
563 TCTCTGGGAAGCTGTACCTCTATGGAATGCAAGCCATGTGTAGCTGTATGATCT 622
541 TCTCTGGGAAGCTGTACCTCTATGGAATGCAAGCCATGTGTAGCTGTATGATCT 600
623 TTGAGATGATTTCTGAGCATGTGTGTGCTGTGGCATCGGAAACAGCTCCGTACTG 682
601 TTGAGATGATTTCTGAGCATGTGTGTGCTGTGGCATCGGAAACAGCTCCGTACTG 660
LT 6
S
3596
NITTON B1653596 777 bp mRNA linear EST 12-SEP-2001
603300230P1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5340823 5'
mRNA sequence.
B1653596
ION B1653596.1 GI:15567832
ORDS
CR EST.
Mus musculus (house mouse)
GANISM
Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 777)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LRAM1865 row: 1 column: 08
High quality sequence stop: 760.
Location/Qualifiers
1. 777
SOURCE
/organism="Mus musculus"
/mol_type="mRNA"
/accession="129_C57BL/6J_PVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5340823"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_1lb="NCI_CGAP_Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
COUNT 168 a 200 c 206 g 203 t
N
TY Match 96.4%; Score 657.4; DB 12; Length 777;
t Local Similarity 99.6%; Pred. No. 1.3e-163;

Matches 660; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 1 AAATACCTGCTCTGCTCTTCAATTGCTCTTCTGCTGCTGAGGCGGTGATCCTAGT 60
DB 10 AAATACCTGCTCTGCTCTTCAATTGCTCTTCTGCTGCTGAGGCGGTGATCCTAGT 69
QY 61 GTAGCTGTGTGCTGCTGATGATCCAGAGACCAAGCTGCTGCTGCTGCTGCTGCTG 120
DB 70 GTAGCTGTGTGCTGCTGATGATCCAGAGACCAAGCTGCTGCTGCTGCTGCTGCTG 129
QY 121 AACAAACCGGCAACCAACCTTCTAGCTGAGCATCTAATTCTCATTTGCTGCTGAGCT 180
DB 130 AACAAACCGGTAACCAACCTTCTAGCTGAGCATCTAATTCTCATTTGCTGCTGAGCT 189
QY 181 GTAGATGATGTTGATGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 190 GTAGATGATG-TTGATGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 248
QY 241 CTGGGAGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 249 CTGGGAGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 308
QY 301 TGGGCTTCTGTAACCAAGACCAAGATGCGCAAGATGGAAGAGTTCTATGACCAAGCC 360
DB 309 TGGGCTTCTGTAACCAAGACCAAGATGCGCAAGATGGAAGAGTTCTATGACCAAGCC 368
QY 361 CTTCAGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
DB 369 CTTCAGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 428
QY 421 CATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 429 CATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 488
QY 481 AGGAACAGCTGTGTCTGCTGAGGCGGCAATGATCTACCCCTTACTGAGCAAGATGTT 540
DB 489 AGGAACAGCTGTGTCTGCTGAGGCGGCAATGATCTACCCCTTACTGAGCAAGATGTT 548
QY 541 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 549 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 608
QY 601 GTGTACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 659
DB 609 GTGTACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668
QY 660 CCGGAACAGCTCCGTGACTGAG 682
DB 669 CCGGAACAGCTCCGTGACTGAG 691
RESULT 7
BO714590 956 bp mRNA linear EST 16-JUL-2002
LOCUS
DEFINITION
AGENCOURT 8292060 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6308902
BO714590
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
MUS
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 956)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
cDNA Library Preparation: Resgen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

Best Local Similarity	98.5%	Pred. No. 2.7e-160;
Matches	651; Conservative	0; Mismatches 10; Indels 0; Gaps 0;
Qy	1 A A A T A C C T G C T C T T G T C T T G A A T T T G C T T T G G C T G G C T G A G G C G T G A T C T A G T	60
Db	160 A A A T A C C T G C T C T T G T C T T C A A T T T G C T T T G G C T G G C T G A G G C G T A T C T A G T	219
Qy	61 G T A G C T C T G T G T G G C G A T G A T T C C A C A G C A C C A G C C T G C T G T A C T G A A C T G G A	120
Db	220 G T A G C T C T G T G T G G C G A T G A T T C C A C A G C A C C A G C C T G C T G T A C T G A A C T G G A	279
Qy	121 A A C A A C C G G A C C C A C A C C T T C T A G C T G G G A C T T A C A T T T C A T T G C T G T G G A G C T	180
Db	280 A A C A A C C G G A C C C A C A C C T T C T A G T G T G G A C A T C A C A T T T C T A C T T G C T G T G G A G C T	339
Qy	181 G T G A T G A T G T T T G A G C G T T C T T G G G G T G C T A T G C A G A G T C C A G T G T C T G	240
Db	340 G T G A T G A T G T T T G A G C G T T C T T G G G G T G C T A T G C A G A G T C C A G T G T C T G	399
Qy	241 C T G G G G A A G T T C T C A C C T G C T T G T A T C C T G T	300
Db	400 C T G G G G A A G T T C T C A C C T G C T T G T A T C C T G T	459
Qy	301 T G G G G C T T C G T A A C A A A G A C A G A T G C C A G G A T G T A A G A G A G T T C T A T G A C C A G G C	360
Db	460 T G G G G C T T C G T A A C A A A G A C A G A T G C C A G G A T G T A A G A G A G T T C T A T G A C C A G G C	519
Qy	361 C T T C A G A C A G C T G T A T G A T A T G A T A T G C C A A C A T C C A A G C T G T G T A A G A C T T T C	420
Db	520 C T T C A G A C A G C T G T A T G A T A T G A T A T G C C A A C A T C C A A G C T G T G T A A G A C T T T C	579
Qy	421 C A T G A G A G C T C A A C T G T T T T T T G T G C T C C A A G C A C T G A C C A C T G A C T A C C A T A C T G	480
Db	580 C A T G A G A G C T C A A C T G T T T T T T G T G C T C C A A G C A C T G A C C A C A C T G A C T A C C A T A C T G	639
Qy	481 A G A A C A G C C T G T G T C C C T C A G G C G G C A C A T A C T A C C C C T T A C T G A G A G A A G A T T G T	540
Db	640 A G A A C A G C C T G T G T C C C T C A G G C G G C A C A T A C T A C C C C T T A C T G A G A G A A G A T T G T	699
Qy	541 C A T C A G A A A A T C A T A G A G C T C T T C T C T G G A A G C T T A C C T C A T T G A A T T G A G C A C A T T	600
Db	700 C A T C A G A A A A T C A T A G A G C T C T T C T C T G G A A G C T T A C C T C A T T G A A T T G A G C A C A T T	759
Qy	601 G T G T A G C T C T C A T T A T G A T C T T T T G A T G A T T C T A G A C A T G T G C T G T G T G C A T C	660
Db	760 G T G T A G C T C T C A T T A T G A T C T T T T G A A G A T T C T C A C A A G T G T G T G T G C G A C	819
Qy	661 C 661	
Db	820 C 820	

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 plate: LLAM1084 row: k column: 01
 High quality sequence stop: 741.

RES
 source
 Location/Qualifiers
 1. 759

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:5030136"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Mam5"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt;
 Site 2: Not; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennigsen/Robin Humphreys,
 NIH"

COUNT 166 a 194 c 209 g 190 t

ry Match 94.0%; Score 641.2; DB 12; Length 759;
 Local Similarity 99.1%; Pred. No. 2.6e-159;
 ches 676; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

1 AATATCTGCTCTTCTGCTCAATTTGCTCTTCTGCTGCTGAGGCTGATCTAGT 60
 13 AATATCTGCTCTTCTGCTCAATTTGCTCTTCTGCTGCTGAGGCTGATCTAGT 72
 61 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 73 GTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 132
 121 AACCAACCGGACCCCAACACCTTCTAGCTGAGGCTGCTGCTGCTGCTGCTGCT 180
 133 AACCAACCGGACCCCAACACCTTCTAGCTGAGGCTGCTGCTGCTGCTGCTGCT 192
 181 GTATGATGTTTGTAGGCTCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 193 GTATGATGTTTGTAGGCTCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 252
 241 CTGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 253 CTGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
 301 TGAGGCTTCTGTAACAAGACAGATGCGCAAGATGTAAGAGCTTCTATGACAGGCC 360
 313 TGAGGCTTCTGTAACAAGACAGATGCGCAAGATGTAAGAGCTTCTATGACAGGCC 372
 361 CTTGACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 373 CTTGACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
 421 CATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 433 CATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
 481 AGGAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 493 AGGAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552
 541 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 553 CATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 610
 601 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 611 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669

Qy 661 CGAAGAGCTCCGTACTAG 682
 Db 670 CGAAGAGCTCCGTACTAG 691

RESULT 11

LOCUS B1695388 759 bp mRNA linear EST 18-SEP-2001
 DEFINITION 603345472P1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5373184 5',
 mRNA sequence.
 ACCESSION B1695388
 VERSION B1695388.1 GI:15658017
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Mammalia; Butcheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 TITLE NIH-MGC <http://img.nci.nih.gov/>
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 plate: LLAM11949 row: p column: 17
 High quality sequence stop: 754.

FEATURES

source

1. 759
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone="IMAGE:5373184"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Mam2"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt;
 Site 2: Not; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 164 a 194 c 204 g 197 t

Query Match 93.8%; Score 640; DB 12; Length 759;
 Best Local Similarity 99.7%; Pred. No. 5.4e-159;

Matches 662; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 18 CTTCAATTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 77
 Db 1 CTTCAATTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 Qy 78 TATGATTCACAGACCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 137
 Db 61 TATGATTCACAGACCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 Qy 138 CACTTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 197
 Db 121 CACTTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 178
 Qy 198 CTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
 Db 179 CTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 238
 Qy 258 CTGCTTGTGATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 317
 Db 239 CTGCTTGTGATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 298

```

/mol_type="RNA"
/db_xref="taxon:10090"
/clone="IMAGE:633531"
/lab_host="DH10B (phage-resistant)"
/clone_11b="N1H_MGC_130"
/note="Organ: oocytes; Vector: PCMV-SPOrt6.1.ccd;
Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 1.95 kb.
Constructed by Resgen, Invitrogen Corp. Note: this is a
N1H MGC library."
COUNT
181 a 265 c 237 g 215 t 5 others

```

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Bukaryote; Metacoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (Pages 1 to 642)				
NIH-MGC	http://mgc.nci.nih.gov/			
National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished				
Contact: Robert Strausberg, Ph.D.				
Email: cgabds-remail.nih.gov				
Tissue Procurement: Jeffrey E. Green, M.D.				
cDNA Library Preparation: Life Technologies, Inc.				
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
DNA Sequencing by: Incyte Genomics, Inc.				
Clone distribution: MGC clone distribution information can be				

Query Match	92.9%	Score 633.4	DB 12	Length 846
Best Local Similarity	99.7%	Pred. No. 3.2e-157		
Matches 645	Conservative 0	Mismatches 1	Indels 1	Gaps 1
QY	36	GCYGCCTGAGGCGCTGATCTCTAGGCTGAGCTCTGTGTGTCCTGATCAAGATCCACAGACAC	95	
DB	17	GCYGCCTGAGGCGCTGATCTCTAGGCTGAGCTCTGTGTGTCCTGATCAAGATCCACAGACAC	76	
QY	96	CAGCCTGCTGTAACCTGGAACCTGGAAAACAACCGGCAACCCACACCTTCTACGTGGGCAT	155	
DB	77	CAGCCTGCTGTAACCTGGAACCTGGAAAACAACCGGCAACCCACACCTTCTACGTGGGCAT	136	
QY	156	CTACATCTCATCTGCTGTGGGAGCTGATGATGTTGATAGGCTTCCTGGGGGTGCTATG	215	
DB	137	CTACATCTCATCTGCTGTGGGAGCTGATGATGTTGATAGGCTTCCTGGGGGTGCTATG	195	
QY	216	GGCCATCCAGAGATCCAGATGTCTGTCTGGGAGCGTTCCTCACCTGCTTGTGATCTGTT	275	
DB	196	GGCCATCCAGAGATCCAGATGTCTGTCTGGGAGCGTTCCTCACCTGCTTGTGATCTGTT	255	
QY	276	TGCCGTGAGAGGTGCTGACAGGCATCTGGGGCTTCGTGAAACAAGACCAATGGCCACAGA	335	
DB	256	TGCCGTGAGAGGTGCTGACAGGCATCTGGGGCTTCGTGAAACAAGACCAATGGCCACAGA	315	
QY	336	TGTGAAAGCATTTCTATGACCAAGGCCCTTCAGCAAGCTGTGATGATGATATGACCAACA	395	
DB	316	TGTGAAAGCATTTCTATGACCAAGGCCCTTCAGCAAGCTGTGATGATGATATGACCAACA	375	
QY	396	TGCCAAGCTGTGTGTAAGCTTTCATGAGACGCTCAACTGTGTGCTCCACGCACT	455	
DB	376	TGCCAAGCTGTGTGTAAGCTTTCATGAGACGCTCAACTGTGTGCTCCACGCACT	435	

[illegible]

TLT 15
 1041
 Bf181041 806 bp mRNA linear EST 31-OCT-2000
 60180508991 NCI CGAP Mam5 Mus musculus cDNA clone IMAGE:4035767 5',
 mRNA sequence.
 Bf181041
 ASTON

ION	BF181041.1	GI:11059183
ORDS	EST.	
CE	Mus musculus (house mouse)	
GANISM	Mus musculus	

RANCE
 THOSE
 TLE
 UNPL
 UNT
 Contact: Robert Straubenberg, ph. D.
 Unpublished
 National Institutes of Health, Mammalian Gene Collection (MGC)
 NIH-MGC <http://mgc.nci.nih.gov/>
 1 (bases 1 to 806)
 Mammalia; Choriata; Craniata; Vertebrata; Euteleostomi;
 Euryptera; Metazoa; Choriata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 806)
 NIH-MGC <http://mgc.nci.nih.gov/>
 Unpublished
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Contact: Robert Straubenberg, ph. D.
 Unpublished

Email: cgsabes@remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E.S. Consortium (LINTL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E.S. Consortium/LINTL at:
<http://image.llnl.gov>
Plate: LHM9310 row: b column: 24
High quality sequence stop: 742.

source	Location/Qualifiers
UBS	1. .806

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="4035767"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam5"
/name="Organ: mammary; Vector: pCMV-SPORE6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lochar Hennighausen/Robin Humphreys, NIH"
COUNT IN 176 a 186 c 224 g 220 t

```

217 March 92.0%; Score 627.2; DB 10; Length 806;
3rd Local Similarity 98.8%; Pred. No. 1.4e-155;
Matches 674; Conservative 0; Mismatches 3; Indels 5; Gaps 4;

1 AATACCTGCTCTTGGTCTTCAATTGCTCTTGGCTGGAGGCGATCTAGT 60
14 AATACCTGCTCTTGGTCTTCAATTGCTCTTGGCTGGAGGCGATCTAGT 73
61 GAGCTGTGTGGTGCATCATCAAGACCAACGCGCTGTCTCTGGAATGGGA 120

Db	74	GTAGCTGTGTGGTGTGATCATCAAGACCAAGCGCTGTGTAACTGTGAAGTGGGA	113
Qy	121	AACTAAACCGGACACCCAAACCTTTCTAAGTGGGATCTAACTTCTATGCTGTGGAGCT	180
Db	134	AACTAAACCGGACACCCAAACCTTTCTAAGTGGGATCTAACTTCTCA-TGCTGTGGAGCT	192
Qy	181	GTATGATATTTGTAGAGCTTCCCTGGGGTGTCTAATGGGGCCATCCAGAGTCCAGTGTCTG	246
Db	193	GTATGATATTTGTAGAGCTTCCCTGGGGTGTCTAATGGGGCCATCCAGAGTCCAGTGTCTG	255
Qy	241	CTGGGACGTTCTTCAACCTGACCTTGTGATCTGTGTGCTGTGAGGTGGCTGACGACATC	300
Db	253	CTGGGACGTTCTTCAACCTGACCTTGTGATCTGTGTGCTGTGAGGTGGCTGACGACATC	312
Qy	301	TGGGGCTTCGTAAACAAAGAACCGATTCGCAAGATGTGAAGCATTTCTATGACAGGCC	366
Db	313	TGGGGCTTCGTAAACAAAGAACCGATTCGCAAGATGTGAAGCATTTCTATGACAGGCC	372
Qy	361	CTTCAGCAAGCTGTGATGTGATATGTGCACAAATGTCCAAAGGCTGTGGTGAAGACTTTC	420
Db	373	CTTCAGCAAGCTGTGATGTGATATGTGCACAAATGTCCAAAGGCTGTGGTGAAGACTTTC	432
Qy	421	CATGAGACGCTCAACTGTGTGTGGCTTCCAAAGCATGACACACTGACCTACCAACCATATCTG	480
Db	433	CATGAGACGCTCAACTGTGTGTGGCTTCCAAAGCATGACACACTGACCTACCAACCATATCTG	492
Qy	481	AGGAACACGCTGTGTGCTCCCTCAGGGGGCAACATCTCACCCCTTACTGTGACGAAGATTGT	540
Db	493	AGGAACACGCTGTGTGCTCCCTCAGGGGG-ACCATATCTCACCCCTTACTGTGACGAAGATTGT	551
Qy	541	CATCAGAAATCGATGAGCTCTTCTGTGGGAAGCTGTAACCTCATTTGGAATGTGACGACATT	600
Db	552	CATCAGAAATCGATGAGCTCTTCTGTGGGAAGCTGTAACCTCATTTGGAATTTG-AGCATTT	609
Qy	601	GTGTGAGCTGTCAATTATGATCTTTGAGATGATTCTGAGCATGATGTGTGCTGTGGCATC	660
Db	610	GTGTGAGCTGTCAATTATGATCTTTGAGATGATTCTGAGCATGATGTGTGCTGTGGCATC	669
Qy	661	CGAAGACGCTCCGTGTACTGAG	682
Db	670	CGAAGACG-TCGATGTACTGAG	690

Search completed: December 9, 2003, 08:29:05
Job time : 2630 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

protein - protein search, using sw model

on: December 9, 2003, 08:29:17 ; Search time 74 Seconds

(without alignments)
484.760 Million cell updates/sec

US-10-035-914-2

ience: 1 KYLLFVFNFWFIAGVILG.....IFEMILSMVLCCGIRNSSVY 226

..ing cable: BLOSUM62

ing cable: BLOSUM62

Gapop 10.0 , Gapext 0.5

1107863 beqs, 158726573 residues

1 number of hits satisfying chosen parameters: 1107863

```
num DB seq length: 0
num DB seq length: 2000000000
```

num DB seq length: 200000000000

-processing: Minimum Match 0%

Listing first 45 summaries

Listing first 45 summaries

base : A_Geneseq_19Jun03: *

1:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1960.DAT
2:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1961.DAT
3:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1962.DAT
4:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1963.DAT
5:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1964.DAT
6:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1965.DAT
7:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1966.DAT
8:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1967.DAT
9:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1968.DAT
10:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1969.DAT
11:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1970.DAT
12:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1991.DAT
13:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1992.DAT
14:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1993.DAT
15:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1994.DAT
16:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1995.DAT
17:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1996.DAT
18:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1997.DAT
19:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1998.DAT
20:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA1999.DAT
21:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA2000.DAT
22:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA2001.DAT
23:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA2002.DAT
24:	/SIDS1.gcgdata/genseeq/genseeq-emb1/AA2003.DAT

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ID	Score	Query Match	Length	DB	Description
1	1177	99.7	258	24	ABG76331
2	1103	93.5	236	22	AA64826
3	1103	93.5	236	23	ABP52845
4	1103	93.5	236	23	ABG78499
5	1103	93.5	236	24	ABU07302
6	1103	93.5	236	24	ABU07303
7	1103	93.5	236	24	ABU07305
8	1103	93.5	236	24	ABU07310
9	1082	91.7	236	20	AA093153

10	1082	91.7	236	24	ABU073307	Human expressed pr
11	1034	87.6	262	21	ABBS31254	Human colon cancer
12	1034	87.6	262	24	ABU073304	Human expressed pr
13	1034	87.6	279	22	AAAG69276	Human secreted pr
14	1034	87.6	279	24	ABU073306	Human expressed pr
15	1003	85.0	222	24	ABG763142	Human expressed pr
16	1003	85.0	222	24	ABG763142	Truncated form of
17	518.5	43.9	227	17	AAAB66834	Mouse recombinant
18	518.5	43.9	227	17	AAAB66834	Human CD9 sequence
19	518.5	43.9	227	23	AAAB78366	Amino acid sequence
20	518.5	43.9	227	23	AAAB78366	Human CD9 antigen
21	518.5	43.9	227	24	ABU05057	Human expressed pr
22	518.5	43.9	228	22	ABU05060	Human expressed pr
23	518.5	43.9	228	22	ABU05060	Human wound healin
24	518.5	43.9	228	24	ABU05048	Human expressed pr
25	518.5	43.9	228	24	ABU05050	Human expressed pr
26	518.5	43.9	228	24	ABU05052	Human expressed pr
27	518.5	43.9	228	24	ABU05053	Human expressed pr
28	518.5	43.9	228	24	ABU05059	Human expressed pr
29	518.5	43.9	275	21	AAAB433936	Human expressed pr
30	518.5	43.9	275	22	AAAG75156	Human cancer assoc
31	518.5	43.9	275	22	ABU05054	Human colon cancer
32	518.5	43.9	275	24	ABU05055	Human expressed pr
33	513.5	43.5	228	13	AAAR27525	Human expressed pr
34	512.5	43.4	228	24	ABU05056	Metastasis control
35	510.5	43.3	226	22	ABU05049	Human expressed pr
36	510.5	43.3	226	23	ABBA44580	Human expressed pr
37	453.5	38.4	221	19	AAAB67234	Mouse wound healin
38	453.5	38.4	221	19	AAAB67234	Mouse ischaemic co
39	453.5	38.4	221	24	ABU69143	Clone HAID05 5'of
40	425	36.0	229	22	AAAG73745	Human NOXV polyep
41	424.5	36.0	221	24	ABG14067	Human colon cancer
42	403	34.2	113	20	AAAB69142	Novel human digno
43	403	34.2	113	24	AAAY09155	Human NOXV polyep
44	401	34.0	116	24	ABU07309	EC2-R186 fragment
45	401	34.0	116	20	AAAY09154	Human expressed pr
				24	ABU07308	EC2 fragment clone
				20	AAAY09154	Human expressed pr

ALIGNMENTS

RESULT 1

ID	ABG76341	standard; Protein; 258 AA.
XY		

AC ABG76341;

DT 10-MAY-2003 (first entry)

Recombinant mouse protein, CD81.

mouse; protein targeting; exosome; lactadherin; C1 domain; C2 domain; membrane vesicles; mutant; subcellular localization

Mus. 8p.

XXXXXX

XX XX

SECRET

XX 17-ATG-2001
PB

20-DEC-2001; 200105-343991P

XX

XX

XX

targeting polypeptides to exosomes providing a chimeric genetic

construct and introducing the construct into exosome-producing cells in vivo or ex vivo

Example 6; Page 87-88; 94pp; English.

The present invention relates to a method and compounds for targeting polypeptides to exosomes. The method comprises providing a chimeric polypeptide construct encoding the polypeptide fused to a targeting polypeptide comprising lactadherin or its portion comprising a functional C1 and/or C2 domain, and introducing the construct into exosome-producing cells in vivo or ex vivo, to generate recombinant exosomes. The method is useful for targeting proteins to membrane vesicles, particularly exosomes, and is useful in experimental research, therapeutic, prophylactic, and diagnostic areas. The present sequence represents a recombinant mouse protein.

Sequence 258 AA;

CY Match 99.7%; Score 1177; DB 24; Length 258;
Local Similarity 99.6%; Pred. No. 5e-127; Indels 0; Gaps 0;
Seqs 225; Conservative 1; Mismatches 0;

```

1 KYLLFVFNFMVFLAGVILGVALLMRHPDPTSLVLELGNKPAPNTFYGIYLAVGA 60
  |||||
11 KYLLFVFNFMVFLAGVILGVALLMRHPDPTSLVLELGNKPAPNTFYGIYLAVGA 70
  |||||
61 VMFVFGFLGCGAIOESQCLGTFPTCLVILFACVAAAGVFNKQDIAVDKQFYDQA 120
  |||||
71 VMFVFGFLGCGAIOESQCLGTFPTCLVILFACVAAAGVFNKQDIAVDKQFYDQA 130
  |||||
121 LQAAVMDDDANNNAKAVKTFHETLNCSSNALTTTLTILNSLCPGSGNITPLAQDC 180
  |||||
131 LQAAVMDDDANNNAKAVKTFHETLNCSSNALTTTLTILNSLCPGSGNITPLAQDC 190
  |||||
181 HOKIDELFSGLKYLIGIAIVAVIMIFEMILSMVLCGIRNSVY 226
  |||||
191 HOKIDELFSGLKYLIGIAIVAVIMIFEMILSMVLCGIRNSVY 236
  |||||

```

1 2
826
AAG64826 standard; protein; 236 AA.

AAG64826;

20-SEP-2001 (first entry)

Chronic hepatitis treatment related protein SEQ ID NO: 7.

Chronic hepatitis; viral antigenic protein; hepatitis C; hepatitis B.

Homo sapiens.

WO200147545-A1.

05-JUL-2001.

28-DEC-2000; 2000WO-JP09393.

28-DEC-1999; 99JP-0374087.

(SUMU) SUMITOMO PHARM CO LTD.

Tohdoh N, Murata M, Enjoji T;

WPI; 2001-425585/45.

Treatment and prevention of chronic hepatitis

Example 1; Page 73-74; 128pp; Japanese.

The present invention describes a method of preventing and treating chronic hepatitis, involving administering an oligopeptide which

CC (a) has binding affinity towards the viral antigenic protein;
CC (b) inhibits binding affinity of the virus towards the receptor protein
CC of the target cell and
CC (c) has analogy with the receptor protein at the amino acid level. This
CC can be used to prevent and treat hepatitis B and C. The present sequence
CC is a protein described in the exemplification of the invention.
XX
SQ Sequence 236 AA;

Query Match 93.5%; Score 1103; DB 22; Length 236;
Best Local Similarity 91.6%; Pred. No. 1.5e-118;
Matches 207; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 KYLLFVFNFMVFLAGVILGVALLMRHPDPTSLVLELGNKPAPNTFYGIYLAVGA 60
  |||||
DB 11 KYLLFVFNFMVFLAGVILGVALLMRHPDPTSLVLELGNKPAPNTFYGIYLAVGA 70
  |||||
QY 61 VMFVFGFLGCGAIOESQCLGTFPTCLVILFACVAAAGVFNKQDIAVDKQFYDQA 120
  |||||
DB 71 VMFVFGFLGCGAIOESQCLGTFPTCLVILFACVAAAGVFNKQDIAVDKQFYDQA 130
  |||||
QY 121 LQAAVMDDDANNNAKAVKTFHETLNCSSNALTTTLTILNSLCPGSGNITPLAQDC 180
  |||||
DB 131 LQAAVMDDDANNNAKAVKTFHETLNCSSNALTTTLTILNSLCPGSGNITPLAQDC 190
  |||||
QY 181 HOKIDELFSGLKYLIGIAIVAVIMIFEMILSMVLCGIRNSVY 226
  |||||
DB 191 HOKIDELFSGLKYLIGIAIVAVIMIFEMILSMVLCGIRNSVY 236
  |||||

```

RESULT 3
ABP52845
ID ABP52845 standard; Protein; 236 AA.

XX ABP52845;
XX
DT 04-NOV-2002 (first entry)

XX Human CD81 protein SEQ ID NO:5.

XX Human; p15 region; chromosome 11; tumour growth; infertility; cytostatic;
XX hepatitis C virus infection; anti-infectivity; gene therapy; CD 81;
XX cluster of differentiation antigen 81.

XX Homo sapiens.

XX WO200261085-A2.

XX 08-AUG-2002.

XX 31-OCT-2001; 2001WO-US45381.

XX 31-OCT-2000; 2000US-244705P.

XX (RYAN/) RYAN J W.

XX Ryan JW;

XX WPI; 2002-619251/66.

XX N-PSDB; ABQ75121.

XX New genes obtainable from the p15 region of human chromosome 11 (e.g.
XX human achaete-scute homolog 2), useful in gene therapy, particularly
XX for preventing or treating tumor growth, infertility or hepatitis C
XX virus infection -
XX
XX Claim 1; Fig 1; 94pp; English.

XX The present invention describes an isolated genomic polynucleotide (I),
XX which is obtainable from the p15 region of human chromosome 11. (HASH2),
XX Specifically described are the human achaete-scute homologue 2 (HASH2),
XX SMS3, tumour suppressing sub-transferable candidate 6 (TSSC6), ribosomal
XX protein L26 (RIBO26), cluster of differentiation antigen 81 (CD 81) and

tumour suppressing sub-transferable candidate 4 (TSSC4) genes as given in ABQ75117 to ABQ75122, encoding the proteins given in ABP52841 to ABP52846. Also described is an isolated polynucleotide obtainable from the p15 region of chromosome 11 having the sequence, which comprises any of two nucleotide sequences given in ABQ75123 and ABQ75124. (1) has cytostatic and anti-fertility activities and can be used in gene therapy. The HNSH2, SMS3, TSSC6, RIBO26, CD 81 or TSSC4 polynucleotide sequences can be used for preventing, treating or ameliorating a medical condition, or for manufacturing a gene therapy for the prevention, treatment or amelioration of such medical conditions. An antisense oligonucleotide or mimetic of (1) can also be useful for manufacturing a conditions. These conditions include tumour growth, infertility, or hepatitis C virus infection.

Sequence 236 AA;

ery Match 93.5%; Score 1103; DB 23; Length 236;
at Local Similarity 91.6%; Pred. No. 1.5e-118;
tches 207; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

```

1 KTLFVFNFMVWLAGVILGVALMLRHPDPTSLYLKLNKRPAPNTFYVGIYILAVGA 60
  |||||
11 KTLFVFNFMVWLAGVILGVALMLRHPDPTSLYLKLNKRPAPNTFYVGIYILAVGA 70
  |||||
61 VMFVFGFGCGAIGSOCLGTFPTCLVILFACVAAAGVNGFVNKQDIADVKQPTDQA 120
  |||||
71 VMFVFGFGCGAIGSOCLGTFPTCLVILFACVAAAGVNGFVNKQDIADVKQPTDQA 130
  |||||
121 LQQAAMDNDANNAKAVVKTFTHTLNCSSNALTTTLTILNSLCPSSGNILTPLODDC 180
  |||||
131 LQQAAMDNDANNAKAVVKTFTHTLNCSSNALTTTLTILNSLCPSSGNILTPLODDC 190
  |||||
181 HOKIDELFSGKLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSYV 226
  |||||
191 HOKIDELFSGKLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSYV 236
  |||||

```

XT 4

1499 AAG78499 standard; protein; 236 AA.

AAG78499;

19-JUN-2002 (first entry)

Human tetraspanin CD81.

Human; CD81; tetraspanin; protein coordinate data.

Homo sapiens.

Key Location/Qualifiers
Domain 113..201
/note="extracellular loop"

MO200202632-A2.

10-JAN-2002.

03-JUL-2001; 2001WO-1B01458.

03-JUL-2000; 2000GB-0016362.

(CHIR-) CHIRON SPA.

Bolognesi M, Grandi G;

WPI; 2002-154731/20.

Modulating (M1) the ability of a tetraspanin protein to form a dimer with the wild-type TTS, by mutating at least one of the amino acids corresponding to specific human CD81 residues

XX Disclosure; Page 2; 56pp; English.

PS This invention relates to the modulation of the ability of a tetraspanin
XX (TTS) protein to form a dimer with the wild-type TTS, comprising
XX mutating at least one of the amino acids corresponding to human CD81
XX residues, 114, 119, 123, 125, 126, 129, 142, 146, 149, 150, 153, 154,
XX 197, 198, 199 and 200. Methods used are useful for modulating the
XX ability of a TTS protein to form a dimer with the wild-type TTS and
XX are useful for identifying a compound that inhibits or promotes TTS
XX dimerization. A computer-based method is also used for identifying a
XX ligand which can interact with TTS to promote or inhibit its ability to
XX dimerise. Certain products from the methods are useful as a
XX pharmaceutical, for treating a patient. This sequence represents
XX human CD81, a tetraspanin.

SQ Sequence 236 AA;

Query Match 93.5%; Score 1103; DB 23; Length 236;
Best Local Similarity 91.6%; Pred. No. 1.5e-118;
Matches 207; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 KTLFVFNFMVWLAGVILGVALMLRHPDPTSLYLKLNKRPAPNTFYVGIYILAVGA 60
  |||||
DB 11 KTLFVFNFMVWLAGVILGVALMLRHPDPTSLYLKLNKRPAPNTFYVGIYILAVGA 70
  |||||
QY 61 VMFVFGFGCGAIGSOCLGTFPTCLVILFACVAAAGVNGFVNKQDIADVKQPTDQA 120
  |||||
DB 71 VMFVFGFGCGAIGSOCLGTFPTCLVILFACVAAAGVNGFVNKQDIADVKQPTDQA 130
  |||||
QY 121 LQQAAMDNDANNAKAVVKTFTHTLNCSSNALTTTLTILNSLCPSSGNILTPLODDC 180
  |||||
DB 131 LQQAAMDNDANNAKAVVKTFTHTLNCSSNALTTTLTILNSLCPSSGNILTPLODDC 190
  |||||
QY 181 HOKIDELFSGKLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSYV 226
  |||||
DB 191 HOKIDELFSGKLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSYV 236
  |||||

```

RESULT 5

ABU07302 ID ABU07302 standard; Protein; 236 AA.

XX AC ABU07302;

DT 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #2003.

XX Translational profiling; expressed protein tag; EPT; kinase;

KW phosphatase; protease; protease inhibitor; transporter;

KW cytoskeletal protein; receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer;

KW gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma;

XX leukemia.

XX Homo sapiens.

XX OS MO200202632-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US09671.

XX PR 28-MAR-2001; 2001US-279495P.

XX PR 21-MAY-2001; 2001US-292544P.

XX PR 08-AUG-2001; 2001US-310801P.

XX PR 01-OCT-2001; 2001US-326370P.

XX PR 04-DEC-2001; 2001US-336780P.

XX PR 20-FEB-2002; 2002US-358985P.

XX PR (ZYCO-) ZYCO INC.

Chicz RM, Tomlinson AJ, Urban RG;
WPI: 2003-040607/03.

New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia

Example 2; SEQ ID No 2003; 134pp; English.

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling.

Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 236 AA;

Match 93.5%; Score 1103; DB 24; Length 236;
Local Similarity 91.6%; Pred. No. 1.5e-118;

Matches 207; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

```
1 KTLFVFNFMVFLAGVYILGVALLMRLHDPQTSLLYLKLNKPAWTFYVGIYLLIIVGA 60
11 KTLFVFNFMVFLAGVYILGVALLMRLHDPQTSLLYLKLNKPAWTFYVGIYLLIIVGA 70
61 VMFVGFGLCYGAIQESQCLGTFFTCVILFACVAAAGIWFVNDQIAKDVQKQFYDQA 120
71 VMFVGFGLCYGAIQESQCLGTFFTCVILFACVAAAGIWFVNDQIAKDVQKQFYDQA 130
121 LQQAAMDDANNAAVAVKTFHETLNCSSNALTTTLTILNSLCPSGSNITPLLODC 180
131 LQQAAMDDANNAAVAVKTFHETLNCSSNALTTTLTILNSLCPSGSNITPLLODC 190
181 HOKIDELFSGKLYLIGIAIYVAIVIMIFEMILSMVLCGGRNSSVY 226
191 HOKIDELFSGKLYLIGIAIYVAIVIMIFEMILSMVLCGGRNSSVY 236
```

6
103
ABU07303 standard; Protein; 236 AA.

ABU07303;

29-JAN-2003 (first entry)

Human expressed protein tag (EPT) #2004.

Translational profiling; expressed protein tag; EPT; kinase; phosphatase; protease; protease inhibitor; transporter; cytoskeletal protein; receptor; transcription factor; cancer; MHC; major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

homo sapiens.

WO200278524-A2.

```
XX 10-OCT-2002.
PD 28-MAR-2002; 2002MO-US09671.
XX 28-MAR-2001; 2001US-279495P.
XX 21-MAY-2001; 2001US-292544P.
XX 08-AUG-2001; 2001US-310801P.
XX 01-OCT-2001; 2001US-326370P.
XX 04-DEC-2001; 2001US-336780P.
XX 20-FEB-2002; 2002US-358985P.
XX (ZYCO-) ZYCO5 INC.
XX Chicz RM, Tomlinson AJ, Urban RG;
PI WPI: 2003-040607/03.
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma
XX or leukemia
XX Example 2; SEQ ID No 2004; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor.
XX The polypeptide is useful as an immunogenic composition for eliciting
XX in a mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to
XX this polypeptide, is useful for treating cancer. The polypeptide is
XX also useful for identifying compounds that binds to a naturally
XX processed class I or class II MHC-binding polypeptide. The polypeptides
XX and polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling.
XX Note: This sequence does not appear in the printed specification but was
XX obtained in electronic format directly from WIPO at
XX ftp://wipo.int/pub/published\_pct\_sequences.
XX
XX Sequence 236 AA;
XX
XX Query Match 93.5%; Score 1103; DB 24; Length 236;
XX Best Local Similarity 91.6%; Pred. No. 1.5e-118;
XX Matches 207; Conservative 14; Mismatches 5; Indels 0; Gaps 0;
OY 1 KTLFVFNFMVFLAGVYILGVALLMRLHDPQTSLLYLKLNKPAWTFYVGIYLLIIVGA 60
DB 11 KTLFVFNFMVFLAGVYILGVALLMRLHDPQTSLLYLKLNKPAWTFYVGIYLLIIVGA 70
OY 61 VMFVGFGLCYGAIQESQCLGTFFTCVILFACVAAAGIWFVNDQIAKDVQKQFYDQA 120
DB 71 VMFVGFGLCYGAIQESQCLGTFFTCVILFACVAAAGIWFVNDQIAKDVQKQFYDQA 130
OY 121 LQQAAMDDANNAAVAVKTFHETLNCSSNALTTTLTILNSLCPSGSNITPLLODC 180
DB 131 LQQAAMDDANNAAVAVKTFHETLNCSSNALTTTLTILNSLCPSGSNITPLLODC 190
OY 181 HOKIDELFSGKLYLIGIAIYVAIVIMIFEMILSMVLCGGRNSSVY 226
DB 191 HOKIDELFSGKLYLIGIAIYVAIVIMIFEMILSMVLCGGRNSSVY 236
XX
XX RESULT 7
XX ID ABU07305 standard; Protein; 236 AA.
XX AC ABU07305;
XX
```


ry Match 93.5%; Score 1103; DB 24; Length 236;
 Local Similarity 91.6%; Pred. No. 1.5e-118;
 ches 207; Conservative 14; Mismatches 5; Indels 0; Gaps 0;
 1 KYLLFVNFVFMVLGGVILGVALLMLRHDPTTSLYLIELGNKPAKPTFVGIYLLAVGA 60
 11 KYLLFVNFVFMVLGGVILGVALLMLRHDPTTSLYLIELGNKPAKPTFVGIYLLAVGA 70
 61 VMFVFGFLGCTGAIQESQCLLGTFTCLVILPACFVAAGIWFVFNKQIADKVKQFYDOA 120
 71 VMFVFGFLGCTGAIQESQCLLGTFTCLVILPACFVAAGIWFVFNKQIADKVKQFYDOA 130
 121 LQAAVMDDDANNAKAVKTFHEHTLNCSSNALTTTLTILNSLCPESGNIILTPLLQDDC 180
 131 LQAAVMDDDANNAKAVKTFHEHTLNCSSNALTTTLTILNSLCPESGNIILTPLLQDDC 190
 181 HOKIDELFSGKLYLIGIAIVAVVIMFEMILSMVLCCGIRNSSV 226
 191 HOKIDELFSGKLYLIGIAIVAVVIMFEMILSMVLCCGIRNSSV 236

153
 AAY09153 standard; Protein; 236 AA.
 AAY09153;

07-JUL-1999 (first entry)

human CD81m protein sequence.

CD81 protein; receptor; Hepatitis C virus; HCV; E2 protein; infection; immunogen; therapy; diagnosis; EC2 region.

homo sapiens.

00918198-AA.

15-APR-1999.

06-OCT-1998; 98MO-IB01628.

03-JUN-1998; 98GB-0013560.

06-OCT-1997; 97GB-0021182.

(CHIR-) CHIRON SPA.

brignani S, Grandi G;

PI; 1999-264018/22.

hepatitis C receptor protein CD81

claim 2; Fig 1; 55pp; English.

The invention relates to the use of the CD81 protein that is identified as a receptor for Hepatitis C virus (HCV). Compounds that specifically bind to CD81 can be used for therapy and diagnosis of HCV. The CD81 protein binds to HCV, preferably to the E2 protein. The CD81 protein can be used to raise antibodies, which can then be used in the diagnosis or therapy of HCV. The antibodies can be used to prevent the virus binding to patient cells and being internalized. The CD81 protein can also be used to treat HCV infection. The CD81 protein, a functional equivalent, or a compound that binds to CD81, can be used in the manufacture of a medicament for the treatment and diagnosis of a HCV infection. The CD81 protein can also be used as a protective immunogen in the control of HCV. The present sequence represents the human CD81 protein (SWISSPROT Accn No: P18582; EMBL/GENBANK Accn No: M33680).
 Note: The EMBL/GENBANK Accn number is indicated wrongly as M33680 in the claims; the correct accession number is M33680 for the human CD81 protein.

SQ Sequence 236 AA;
 Query Match 91.7%; Score 1082; DB 20; Length 236;
 Best Local Similarity 90.7%; Pred. No. 4e-116;
 Matches 204; Conservative 14; Mismatches 7; Indels 0; Gaps 0;
 1 KYLLFVNFVFMVLGGVILGVALLMLRHDPTTSLYLIELGNKPAKPTFVGIYLLAVGA 60
 11 KYLLFVNFVFMVLGGVILGVALLMLRHDPTTSLYLIELGNKPAKPTFVGIYLLAVGA 70
 61 VMFVFGFLGCTGAIQESQCLLGTFTCLVILPACFVAAGIWFVFNKQIADKVKQFYDOA 120
 71 VMFVFGFLGCTGAIQESQCLLGTFTCLVILPACFVAAGIWFVFNKQIADKVKQFYDOA 130
 121 LQAAVMDDDANNAKAVKTFHEHTLNCSSNALTTTLTILNSLCPESGNIILTPLLQDDC 180
 131 LQAAVMDDDANNAKAVKTFHEHTLNCSSNALTTTLTILNSLCPESGNIILTPLLQDDC 190
 181 HOKIDELFSGKLYLIGIAIVAVVIMFEMILSMVLCCGIRNSSV 225
 191 HOKIDELFSGKLYLIGIAIVAVVIMFEMILSMVLCCGIRNSSV 235

RESULT 10

ABU07307 standard; Protein; 236 AA.

ABU07307;

29-JAN-2003 (first entry)

Human expressed protein tag (BPT). #2008.

Translational profiling; expressed protein tag; BPT; kinase;
 phosphatase; protease; protease inhibitor; transporter;
 cytoskeletal protein; receptor; transcription factor; cancer; MEC;
 major histocompatibility complex; myeloma; colon cancer;
 gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma;
 leukemia.

homo sapiens.

WO200278524-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-US09671.

28-MAR-2001; 2001US-279495P.

21-MAY-2001; 2001US-292544P.

08-AUG-2001; 2001US-310801P.

01-OCT-2001; 2001US-326370P.

04-DEC-2001; 2001US-336780P.

20-FEB-2002; 2002US-358985P.

(ZYCO-) ZYCO INC.

Chiez RM, Tomlinson AJ, Urban RG;

WPI; 2003-040607/03.

New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia

Example 2: SEQ ID No 2008; 134pp; English.

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified

polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that bind to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling.

Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 236 AA;

Query Match 91.7%; Score 1082; DB 24; Length 236;
Local Similarity 90.7%; Pred. No. 4e-116;
Conservative 14; Mismatches 7; Indels 0; Gaps 0;

```

1 KTLFVFNFFVFLAGVILGVALMTLRHDPQTSLYLRLGNKRPAPNTFYVGIYILAVGA 60
11 KTLFVFNFFVFLAGVILGVALMTLRHDPQTSLYLRLGNKRPAPNTFYVGIYILAVGA 70
61 VMFVGLFAGCTGAIQESQCLGTFFTCVILFACVAVAGVFNKQDIADVKQFYDQA 120
71 VMFVGLFAGCTGAIQESQCLGTFFTCVILFACVAVAGVFNKQDIADVKQFYDQA 130
121 LQAAVDDDDANNAKAVVTFHETLNCSSNALTTTLTLNSLCPSGNITLPLAQDC 180
131 LQAAVDDDDANNAKAVVTFHETLNCSSNALTTTLTLNSLCPSGNITLPLAQDC 190
181 HOKIDELPSGKLYLIGIAIVAVAVIMIFEMILSMVLCGIRNSV 225
191 HOKIDELPSGKLYLIGIAIVAVAVIMIFEMILSMVLCGIRNSV 235
AAB53254 standard; Protein; 262 AA.

```

AAB53254;

09-MAR-2001 (first entry)

Human colon cancer antigen protein sequence SEQ ID NO:794.

Human; colon cancer; colon cancer antigen; diagnosis; detection; identification; cytostatic; cardioactive; neuroprotective; vulnary; immunomodulatory; muscular; gynaecological; gastrointestinal; nephrotropic; antiinfective; antibacterial; gene therapy; wound; neural disorder; immune system disorder; muscular disorder; reproductive disorder; gastrointestinal disorder; renal disorder; infectious disease; cardiovascular disorder.

Homo sapiens.

W020005531-A1.

21-SEP-2000.

08-MAR-2000; 2000WO-US05883.

12-MAR-1999; 99US-0124270.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2000-587534/55.

N-PSDB; AAC98011.

Colon cancer associated gene sequences, referred to as colon cancer

PT antigens, useful for the treatment, prevention, and diagnosis of colon disorders such as colon cancer -

PS Claim 11; Page 1350-1351; 2104pp; English.

AAC97991 to AAC98763 encode the human colon cancer associated proteins, called human colon cancer antigens, given in AAB53234 to AAB54006. The human colon cancer antigens can have cytostatic, cardioactive, muscular, neuroprotective, immunomodulatory, gynaecological, gastrointestinal, vulnary, nephrotropic, antiinfective and antibacterial activities, and can be used in gene therapy. The colon cancer antigen polynucleotides, proteins and antibodies to the proteins are useful for the prevention, treatment and diagnosis of colon disorders, such as colon cancer. The polynucleotides may be used in diagnostics and research, such as for chromosome identification, and as hybridisation probes. The proteins may also be used to prevent diseases such as neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, wounds, renal disorders, infectious diseases, and cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent sequences used in the exemplification of the present invention.

Sequence 262 AA;

Query Match 87.6%; Score 1034; DB 21; Length 262;
Local Similarity 91.1%; Pred. No. 1.6e-110;
Conservative 14; Mismatches 5; Indels 0; Gaps 0;

```

13 LAGVILGVALMTLRHDPQTSLYLRLGNKRPAPNTFYVGIYILAVGAVMFFGLCTG 72
49 LAGVILGVALMTLRHDPQTSLYLRLGNKRPAPNTFYVGIYILAVGAVMFFGLCTG 108
73 AIGRSQCLGTFFTCVILFACVAVAGVFNKQDIADVKQFYDQAQAAVDDDDAN 132
109 AIGRSQCLGTFFTCVILFACVAVAGVFNKQDIADVKQFYDQAQAAVDDDDAN 166
133 AKAVVTFHETLNCSSNALTTTLTLNSLCPSGNITLPLAQDCGKIDELPSGKL 192
169 AKAVVTFHETLNCSSNALTTTLTLNSLCPSGNITLPLAQDCGKIDELPSGKL 228
193 YLIGIAIVAVAVIMIFEMILSMVLCGIRNSV 226
229 YLIGIAIVAVAVIMIFEMILSMVLCGIRNSV 262

```

Db

ABU07304 standard; Protein; 262 AA.

ABU07304;

29-JAN-2003 (first entry)

Human expressed protein tag (EPT) #2005.

Translational profiling; expressed protein tag; EPT; kinase; phosphatase; protease; protease inhibitor; transporter; cytoskeletal protein; receptor; transcription factor; cancer; MHC; major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

Homo sapiens.

W0200278524-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-US09671.

28-MAR-2001; 2001US-279495P.

21-MAY-2001; 2001US-292544P.

08-AUG-2001; 2001US-310801P.

01-OCT-2001; 2001US-326370P.
04-DEC-2001; 2001US-336780P.
20-FEB-2002; 2002US-358985P.
(ZYCO-) ZYCOs INC.

Chicz RM, Tomlinson AJ, Urban RG;
WPI; 2003-040607/03.

New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia

Example 2; SEQ ID No 2005; 134pp; English.

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing lymphoma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling.

Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 262 AA;

cy Match 87.6%; Score 1034; DB 24; Length 262;
Local Similarity 91.1%; Pred. No. 1.6e-110;
Matches 195; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

```
13 LAGVILGVALLMLRHPDPTSLLYELGNKPAEPNTFYGIYILAVGAVMMFVGLGCG 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
49 LAGVILGVALLMLRHPDPTSLLYELGNKPAEPNTFYGIYILAVGAVMMFVGLGCG 108
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
73 AIOESQCLIGTFEFTCLVILFACVNAAGIMGFVNKQDIADVKQFYDQALQAVVDDANN 132
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
109 AIOESQCLIGTFEFTCLVILFACVNAAGIMGFVNKQDIADVKQFYDQALQAVVDDANN 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
133 AKAVVTFHEHTLCCGSSNALTTTLTILRNSLCPSSGNITLPLLODCHOKIDLPFGSKL 192
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
169 AKAVVTFHEHTLCCGSSNALTTTLTILRNSLCPSSGNITLPLLODCHOKIDLPFGSKL 228
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
193 YLIGIAIIVAVVIMIFEMILSMVLCGGIRNSSY 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
229 YLIGIAIIVAVVIMIFEMILSMVLCGGIRNSSY 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

13
276
YAG89276 standard; Protein; 279 AA.
YAG89276;

11-SEP-2001 (first entry)

human secreted protein, SEQ ID NO: 396.

human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
JENSET.

homo sapiens.

XX
PN WO200142451-A2.
XX
XX 14-JUN-2001.
PD
XX
XX 07-DEC-2000; 2000WO-1B01938.
PP
XX
XX 08-DEC-1999; 99US-0169629.
PR
XX 06-MAR-2000; 2000US-0187470.
XX
XX (GENT) GENSET.
PA
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI
XX WPI; 2001-367870/38.
DR
XX N-PSDB; AAH64879.

PT Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases -
PT
XX
PS Claim 21; Page 881; 921pp; English.

XX
XX The invention relates to full length GENSET human nucleic acids encoding CC
XX potentially secreted proteins. The nucleic acids and the polypeptides CC
XX they encode may be used in the prevention, treatment and diagnosis of CC
XX diseases associated with inappropriate GENSET gene expression. For CC
XX example, they be used to treat disorders associated with decreased CC
XX GENSET gene expression by rectifying mutations or deletions in a CC
XX patient's genome that affect the activity of GENSET or by supplementing CC
XX the patient's own production of GENSET polypeptides. Conversely, CC
XX antisense nucleic acid molecules may be administered to down regulate CC
XX GENSET expression by binding with the cells' own genes and preventing CC
XX their expression. The sense and antisense nucleic acids may also be CC
XX used as DNA probes in diagnostic assays to detect and quantitate the CC
XX presence of similar nucleic acid sequences in samples, and hence to CC
XX determine which patients may be in need of restorative therapy. CC
XX The GENSET polypeptides may be used as antigens in the production of CC
XX antibodies and in assays to identify modulators (agonists and CC
XX antagonists) of GENSET polypeptide expression and activity. The CC
XX present sequence is a GENSET polypeptide of the invention.

Sequence 279 AA;

Query Match 87.6%; Score 1034; DB 22; Length 279;
Best Local Similarity 91.1%; Pred. No. 1.7e-110;
Matches 195; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

```
QY 13 LAGVILGVALLMLRHPDPTSLLYELGNKPAEPNTFYGIYILAVGAVMMFVGLGCG 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 66 LAGVILGVALLMLRHPDPTSLLYELGNKPAEPNTFYGIYILAVGAVMMFVGLGCG 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 73 AIOESQCLIGTFEFTCLVILFACVNAAGIMGFVNKQDIADVKQFYDQALQAVVDDANN 132
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 126 AIOESQCLIGTFEFTCLVILFACVNAAGIMGFVNKQDIADVKQFYDQALQAVVDDANN 185
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 133 AKAVVTFHEHTLCCGSSNALTTTLTILRNSLCPSSGNITLPLLODCHOKIDLPFGSKL 192
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 186 AKAVVTFHEHTLCCGSSNALTTTLTILRNSLCPSSGNITLPLLODCHOKIDLPFGSKL 245
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 193 YLIGIAIIVAVVIMIFEMILSMVLCGGIRNSSY 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 246 YLIGIAIIVAVVIMIFEMILSMVLCGGIRNSSY 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

RESULT 14
ABU07306
ID ABU07306 standard; Protein; 279 AA.
XX
XX ABU07306;
AC
XX
XX 29-JAN-2003 (first entry)
DT
XX

Human expressed protein tag (EPT) #2007.

Translational profiling; expressed protein tag; EPT; kinase; phosphatase; protease; protease inhibitor; transporter; cytoskeletal protein; receptor; transcription factor; cancer; MHC; major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

Homo sapiens.

W0200278524-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-US09671.

28-MAR-2001; 2001US-279495P.

21-MAY-2001; 2001US-292544P.

08-AUG-2001; 2001US-310801P.

01-OCT-2001; 2001US-326370P.

04-DEC-2001; 2001US-336780P.

20-FEB-2002; 2002US-358985P.

(ZYCO-) ZYCOS INC.

Chicz RM, Tomlinson AJ, Urban RG;

WPI; 2003-040607/03.

New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia

Example 2; SEQ ID No 2007; 134pp; English.

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling.

Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 279 AA;

Query Match 87.6%; Score 1034; DB 24; Length 279;
 Local Similarity 91.1%; Pred. No. 1.7e-110;
 Matches 195; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

13 LAGSVILGVALMLRHDPTTSLVLEIGNKRPAPRTFYGYIYLLAVGAVMMFVGLCTG 72
 66 LAGSVILGVALMLRHDPTTSLVLEIGNKRPAPRTFYGYIYLLAVGAVMMFVGLCTG 125
 73 AIOESOCILGFTFCLVILFACVVAAGIWFVNQDAKQVQKPYDQALQOAVDDANN 132
 126 AIOESOCILGFTFCLVILFACVVAAGIWFVNQDAKQVQKPYDQALQOAVDDANN 185
 133 AKAVVKTFFHETLNCSSNALTTTLTILRNSLCPSGGINLTPILQODCHQKIDELPSGKL 192
 186 AKAVVKTFFHETLNCSSNALTTTLTILRNSLCPSGGINLTPILQODCHQKIDELPSGKL 245

QY 193 YLIGIAIYVAIVIMIFEMILSMVLCGIRNSSVY 226
 DB 246 YLIGIAIYVAIVIMIFEMILSMVLCGIRNSSVY 279

RESULT 15

ABG76342
 ID ABG76342 standard; Protein; 222 AA.

AC ABG76342;

DT 10-MAY-2003 (first entry)

DE Truncated form of recombinant mouse protein, CD81.

KW Mouse; protein targeting; exosome; lactadherin; C1 domain; C2 domain; membrane vesicle; mutant; mutein.

OS Mus sp.

PN WO2003016522-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-EP09108.

PR 17-AUG-2001; 2001US-313159P.

PR 26-DEC-2001; 2001US-343991P.

PA (ANOS-) ANOSYS INC.

PI Delcayre A, Le Pecq J;

DR WPI; 2003-268331/26.

PT Targeting polypeptides to exosomes providing a chimeric genetic construct and introducing the construct into exosome-producing cells in vivo or ex vivo

PS Example 6; Page 89; 94pp; English.

CC The present invention relates to a method and compounds for targeting polypeptides to exosomes. The method comprises providing a chimeric genetic construct encoding the polypeptide fused to a targeting polypeptide comprising lactadherin or its portion comprising a functional C1 and/or C2 domain, and introducing the construct into exosome-producing cells in vivo or ex vivo, to generate recombinant vesicles. The method is useful for targeting proteins to membrane vesicles, particularly exosomes, and is useful in experimental research, therapeutic, prophylactic, and diagnostic areas. The present sequence represents the truncated form of a recombinant mouse protein.

Sequence 222 AA;

Query Match 85.0%; Score 1003; DB 24; Length 222;
 Best Local Similarity 99.5%; Pred. No. 4.7e-107;
 Matches 189; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYLLFVNFVFLAGVILGVALMLRHDPTTSLVLEIGNKRPAPRTFYGYIYLLAVGA 60
 DB 11 KYLLFVNFVFLAGVILGVALMLRHDPTTSLVLEIGNKRPAPRTFYGYIYLLAVGA 70
 QY 61 VMMFVGLGCGYGAIOESOCILGFTFCLVILFACVVAAGIWFVNQDAKQVQKPYDQ 120
 DB 71 VMMFVGLGCGYGAIOESOCILGFTFCLVILFACVVAAGIWFVNQDAKQVQKPYDQ 130
 QY 121 LQOAVDDDDANNAKAVVKTFFHETLNCSSNALTTTLTILRNSLCPSGGINLTPILQODC 180
 DB 131 LQOAVDDDDANNAKAVVKTFFHETLNCSSNALTTTLTILRNSLCPSGGINLTPILQODC 190

181 HOKIDELFSG 190
| | | | |
191 HOKIDELFSG 200

h completed: December 9, 2003, 10:29:41
ime : 77 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

on: December 9, 2003, 10:28:23 / Search time 28 Seconds
(without alignments)
341.509 Million cell updates/sec

le: US-10-035-914-2

lect score: 1180
1 KYLLFVNFVFWLAGVILG.....IFEMILSMVLCGIRNSSVY 226

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 328717 seqs, 4231058 residues

1 number of hits satisfying chosen parameters: 328717

num DB seq length: 0
num DB seq length: 2000000000

-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

base :

1: /cgnt2_6/prodata/2/iaa/5A_COMB.pep:*
2: /cgnt2_6/prodata/2/iaa/5B_COMB.pep:*
3: /cgnt2_6/prodata/2/iaa/5A_COMB.pep:*
4: /cgnt2_6/prodata/2/iaa/5B_COMB.pep:*
5: /cgnt2_6/prodata/2/iaa/PCTUS_COMB.pep:*
6: /cgnt2_6/prodata/2/iaa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

o.	Score	Query Match	Length	DB ID	Description
1	518.5	43.9	227	1	US-08-254-493-1
2	518.5	43.9	227	2	US-08-253-751-6
3	518.5	43.9	227	2	US-08-453-925-6
4	518.5	43.9	227	4	US-08-403-253A-6
5	518.5	43.9	227	4	US-08-435-816A-6
6	518.5	43.9	228	1	US-08-408-222B-1
7	298	25.3	237	3	US-08-808-148-3
8	297.5	25.2	219	2	US-08-855-140-4
9	272.5	23.1	267	3	US-08-430-225A-20
10	265.5	22.5	219	2	US-08-855-140-3
11	265.5	22.5	219	2	US-08-807-044-3
12	265.5	22.5	219	5	PCT-US91-04866-2
13	248.5	21.1	253	3	US-09-333-599-4
14	248.5	21.1	253	4	US-09-499-781-4
15	246.5	20.9	236	3	US-08-705-771-19
16	243.5	20.6	253	3	US-09-333-599-2
17	243.5	20.6	253	4	US-09-499-781-2
18	238.5	20.2	280	2	US-08-855-140-1
19	236	20.0	265	2	US-08-807-044-1
20	220.5	18.7	241	3	US-08-808-148-1
21	220.5	18.7	241	3	US-09-020-956-114
22	220.5	18.7	241	3	US-09-030-607-114
23	220.5	18.7	241	4	US-09-439-313-114
24	220.5	18.7	241	4	US-09-352-616A-114
25	220.5	18.7	241	4	US-09-232-149A-114
26	191.5	16.2	252	3	US-08-705-771-17
27	187.5	15.9	281	3	US-08-808-148-4

28	181.5	15.4	245	4	US-09-482-273-133	Sequence 133, App
29	159	13.5	101	3	US-08-905-223-443	Sequence 443, App
30	113.5	9.6	204	4	US-09-149-476-429	Sequence 429, App
31	105	8.9	260	3	US-08-957-130-15	Sequence 15, App
32	100	8.5	258	3	US-08-957-130-13	Sequence 13, App
33	87	7.4	362	4	US-09-252-991A-27993	Sequence 27993, A
34	86	7.3	79	3	US-08-630-172-8	Sequence 8, Appl1
35	86	7.3	79	3	US-09-375-419-8	Sequence 8, Appl1
36	83	7.0	333	4	US-09-170-496D-8	Sequence 8, Appl1
37	83	7.0	333	4	US-09-170-496D-168	Sequence 168, App
38	83	7.0	403	4	US-09-328-352-7791	Sequence 7791, App
39	77.5	6.6	362	3	US-08-513-974B-374	Sequence 374, App
40	77	6.5	60	3	US-09-188-930-181	Sequence 181, App
41	77	6.5	60	3	US-09-188-930-181	Sequence 320, App
42	77	6.5	60	4	US-09-312-283C-181	Sequence 181, App
43	77	6.5	60	4	US-09-312-283C-320	Sequence 320, App
44	77	6.5	505	4	US-09-328-352-7155	Sequence 7155, App
45	77	6.5	993	1	US-07-977-451-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-08-254-493-1
Sequence 1, Application US/08254493
Patent No. 5439886
GENERAL INFORMATION:
APPLICANT: IKEMAYA, SHUTCHI
APPLICANT: KOYAMA, MASARU
APPLICANT: MIYAKE, MASAYUKI
APPLICANT: SENO, MASARU
TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 110 WATER STREET
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/254,493
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/865552
FILING DATE: 09-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 079996-1991
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 085396-1991
FILING DATE: 17-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 022321-1992
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, DAVID S.
REGISTRATION NUMBER: 34235
REFERENCE/DOCKET NUMBER: 41777
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STR UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 227 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: protein
-254-493-1

RY Match 43.9%; Score 518.5; DB 1; Length 227;
c Local Similarity 44.9%; Pred. No. 6.8e-52;
ches 101; Conservative 39; Mismatches 72; Indels 13; Gaps 5;

```
1 KYLLFVNFVFWLGGVILGVALMLRHPDQTSLLYLELNKRPAPNTFYGIYLLAVGA 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
10 KYLLFGNFIFWLAGVIALGVALRPSQTKSIFQETNN--NSFTYGVYLLIGAGA 67
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 VMFVFGTGYGAIQESQCLGTFPTCLVILFACVAAIGFVNNQDIADVKQPTDQA 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 LMMVGFELGCCGAVQESQCLGTFPTCLVILFALIAAIGVSHKDEVIKVQEFYKDT 127
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 LQAVVMDNANNAKAVVTFHEFTLNCCGSNALTTLTTLIR--NSLCPSSGNILTPLLQ 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 YNKKTDEPO--RETLKAHYALNCCG-----LAGGVQFISDICKP-KDVLFTTVK 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

179 DCHQKIDELFSGKLYLIGIAIVAVIMIFEMILSMVLCGIRNS 223
179 SCPDAKEVFENKFIIGAVGIGIAVVMIFGMIFSMILCAIRN 223

253-751-6
nence 6, Application US/08253751
ent No. 5858358
GENERAL INFORMATION:

APPLICANT: June, Carl H.
APPLICANT: Thompson, Craig B.
APPLICANT: Nabel, Gary J.
APPLICANT: Gray, Gary S.
APPLICANT: Remmert, Paul D.
APPLICANT: Freeman, Gordon J.
TITLE OF INVENTION: METHODS FOR SELECTIVELY STIMULATING
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/253,751
FILING DATE: 3 JUNE 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/073,223
FILING DATE: 4 JUNE 1993
APPLICATION NUMBER: US 08/200,947
FILING DATE: 23 FEB 1994
APPLICATION NUMBER: US 07/864,805
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 08/247,505
FILING DATE: 23 MAY 1994
APPLICATION NUMBER: US 07/864,866
FILING DATE: 7 APR 1992
APPLICATION NUMBER: 08/218,155
FILING DATE: 25 MAR 1994
APPLICATION NUMBER: US 07/864,807
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 07/902,467

FILING DATE: 16 JUNE 1992
APPLICATION NUMBER: US 07/275,433
FILING DATE: 23 NOV 1988
ATTORNEY/AGENT INFORMATION:
NAME: MANDRAGOURAS, AMY E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPI-002CPB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-08-253-751-6

Query Match 43.9%; Score 518.5; DB 2; Length 227;
Best Local Similarity 44.9%; Pred. No. 6.8e-52;
Matches 101; Conservative 39; Mismatches 72; Indels 13; Gaps 5;

```
QY 1 KYLLFVNFVFWLGGVILGVALMLRHPDQTSLLYLELNKRPAPNTFYGIYLLAVGA 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10 KYLLFGNFIFWLAGVIALGVALRPSQTKSIFQETNN--NSFTYGVYLLIGAGA 67
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 VMFVFGTGYGAIQESQCLGTFPTCLVILFACVAAIGFVNNQDIADVKQPTDQA 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 68 LMMVGFELGCCGAVQESQCLGTFPTCLVILFALIAAIGVSHKDEVIKVQEFYKDT 127
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 LQAVVMDNANNAKAVVTFHEFTLNCCGSNALTTLTTLIR--NSLCPSSGNILTPLLQ 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 128 YNKKTDEPO--RETLKAHYALNCCG-----LAGGVQFISDICKP-KDVLFTTVK 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 179 DCHQKIDELFSGKLYLIGIAIVAVIMIFEMILSMVLCGIRNS 223
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 179 SCPDAKEVFENKFIIGAVGIGIAVVMIFGMIFSMILCAIRN 223
```

RESULT 3
US-08-453-925-6
Sequence 6, Application US/08453925
Patent No. 5883223
GENERAL INFORMATION:

APPLICANT: June, Carl H., Thompson, Craig B., Nabel, Gary J.
APPLICANT: Gray, Gary S., Remmert, Paul D., Freeman, Gordon J.
TITLE OF INVENTION: METHODS FOR SELECTIVELY STIMULATING
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/453,925
FILING DATE: 30 MAY 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/253,751
FILING DATE: 3 JUNE 1994
APPLICATION NUMBER: US 08/073,223
FILING DATE: 4 JUNE 1993
APPLICATION NUMBER: US 08/200,947
FILING DATE: 23 FEB 1994
APPLICATION NUMBER: US 07/864,805
FILING DATE: 7 APR 1992

APPLICATION NUMBER: US 08/247,505
FILING DATE: 23 MAY 1994
APPLICATION NUMBER: US 07/864,866
FILING DATE: 7 APR 1992
APPLICATION NUMBER: 08/218,155
FILING DATE: 25 MAR 1994
APPLICATION NUMBER: US 07/864,807
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 07/902,467
FILING DATE: 16 JUNE 1992
APPLICATION NUMBER: US 07/275,433
FILING DATE: 23 NOV 1988
ATTORNEY/AGENT INFORMATION:
NAME: MANDRAGOURAS, AMY E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPI-002CPB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
1-453-925-6

Query Match 43.9%; Score 518.5; DB 2; Length 227;
Local Similarity 44.9%; Pred. No. 6.8e-52;
Conservative 39; Mismatches 72; Indels 13; Gaps 5;
1 KYLLFVFNFMVLGAVILGVALMLRHPDPTSLVLELGNKRAPNTEVGYILLAVGA 60
10 KYLLFVFNFMVLGAVILGVALMLRHPDPTSLVLELGNKRAPNTEVGYILLAVGA 67
61 VMNFGFLGCTGALIOESQCLGTFPTCLVILFACGEVAGIWFVNRQIAKDVQFYDQA 120
68 LMLLVGFLGCGGAVQESQCLGTFPTCLVILFACGEVAGIWFVNRQIAKDVQFYDQA 127
121 LQAVVDDNANNAKAVKTFHETLNCSSNALTTLTTLR--NSLCPSGGNITPLDQ 178
128 YNKLTQDPO--RETLKAHYALNCCG-----LAGVQFISDIPK-KVLEFTVVK 178
179 DGHOKIDELFSGLYLIGIAIVAVIMFEMILSWVLCGGIRNS 223
179 SCDDAIKEVFNKPHIIGAVGIGIAVIMFEMILSWVLCGGIRNS 223

4
1-403-253A-6
ence 6, Application US/08403253A
ent No. 6352694
GENERAL INFORMATION:
APPLICANT: June, Carl H., Thompson, Craig B., Nabel, Gary J.
APPLICANT: Gray, Gary S., Remmert, Paul D.
TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,253A
FILING DATE: March 10, 1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/253,964
FILING DATE: 3 JUNE 1994
APPLICATION NUMBER: US 08/073,223
FILING DATE: 4 JUNE 1993
APPLICATION NUMBER: US 08/200,947
FILING DATE: 23 FEB 1994
APPLICATION NUMBER: US 07/864,805
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 08/247,505
FILING DATE: 23 MAY 1994
APPLICATION NUMBER: US 07/864,866
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 08/218,155
FILING DATE: 25 MAR 1994
APPLICATION NUMBER: US 07/864,807
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 07/902,467
FILING DATE: 16 JUNE 1992
APPLICATION NUMBER: US 07/275,433
FILING DATE: 23 NOV 1988
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPI-002CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-403-253A-6

Query Match 43.9%; Score 518.5; DB 4; Length 227;
Best Local Similarity 44.9%; Pred. No. 6.8e-52;
Matches 101; Conservative 39; Mismatches 72; Indels 13; Gaps 5;
1 KYLLFVFNFMVLGAVILGVALMLRHPDPTSLVLELGNKRAPNTEVGYILLAVGA 60
10 KYLLFVFNFMVLGAVILGVALMLRHPDPTSLVLELGNKRAPNTEVGYILLAVGA 67
61 VMNFGFLGCTGALIOESQCLGTFPTCLVILFACGEVAGIWFVNRQIAKDVQFYDQA 120
68 LMLLVGFLGCGGAVQESQCLGTFPTCLVILFACGEVAGIWFVNRQIAKDVQFYDQA 127
121 LQAVVDDNANNAKAVKTFHETLNCSSNALTTLTTLR--NSLCPSGGNITPLDQ 178
128 YNKLTQDPO--RETLKAHYALNCCG-----LAGVQFISDIPK-KVLEFTVVK 178
179 DGHOKIDELFSGLYLIGIAIVAVIMFEMILSWVLCGGIRNS 223
179 SCDDAIKEVFNKPHIIGAVGIGIAVIMFEMILSWVLCGGIRNS 223

RESULT 5
US-08-435-816A-6
Sequence 6, Application US/08435816A
Patent No. 6534055
GENERAL INFORMATION:
APPLICANT: June, Carl H.
APPLICANT: Thompson, Craig B.
APPLICANT: Nabel, Gary J.
APPLICANT: Gray, Gary S.
APPLICANT: Remmert, Paul D.
TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston

```

STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,816A
FILING DATE: May 4, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,253
FILING DATE: 10 MARCH 1995
APPLICATION NUMBER: US 08/253,964
FILING DATE: 3 JUNE 1994
APPLICATION NUMBER: US 08/073,223
FILING DATE: 4 JUNE 1993
APPLICATION NUMBER: US 08/200,947
FILING DATE: 23 FEB 1994
APPLICATION NUMBER: US 07/864,805
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 08/247,505
FILING DATE: 23 MAY 1994
APPLICATION NUMBER: US 07/864,866
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 08/218,155
FILING DATE: 25 MAR 1994
APPLICATION NUMBER: US 07/864,807
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 07/902,467
FILING DATE: 16 JUNE 1992
APPLICATION NUMBER: US 07/275,433
FILING DATE: 23 NOV 1988
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy B.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPT-002CP3
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
435-816A-6

Y Match 43.9%; Score 518.5; DB 4; Length 227;
Local Similarity 44.9%; Pred. No. 6.8e-52;
hes 101; Conservative 39; Mismatches 72; Indels 13; Gaps 5

1 KYTLFVFNFMVPLAGVILGVALMTLRHPDQTSLSLYLELNKRAPIPTFYGLIYLAVGA 60
  KYTLFVFNFMVPLAGVILGVALMTLRHPDQTSLSLYLELNKRAPIPTFYGLIYLAVGA 67
10 KYTLFVFNFMVPLAGVILGVALMTLRHPDQTSLSLYLELNKRAPIPTFYGLIYLAVGA 67
61 VMVVFGLFCYGLNIOSSQCLTFTFTCLVILFACBVAAGIGVFNKQIALKDVCKPDDA 120
  VMVVFGLFCYGLNIOSSQCLTFTFTCLVILFACBVAAGIGVFNKQIALKDVCKPDDA 127
68 LMLVLVGLFCYGLNIOSSQCLTFTFTCLVILFACBVAAGIGVFNKQIALKDVCKPDDA 127
  LMLVLVGLFCYGLNIOSSQCLTFTFTCLVILFACBVAAGIGVFNKQIALKDVCKPDDA 127
121 LQAAVMDDDANNAKAVVKTFRHTLNCSSNALITVLTITLR--NSLCPSSGNILPPLQQ 178
  LQAAVMDDDANNAKAVVKTFRHTLNCSSNALITVLTITLR--NSLCPSSGNILPPLQQ 178
128 YNDLTKTKDPEQ--RETLAIHYALNCCG-----LQGVROPTSDICPK-KVLEFTVK 178
  YNDLTKTKDPEQ--RETLAIHYALNCCG-----LQGVROPTSDICPK-KVLEFTVK 178
179 DCHQKIDELFSGLYILGIAIVAVVIMIFEMILSNVLCGIGINS 223
  DCHQKIDELFSGLYILGIAIVAVVIMIFEMILSNVLCGIGINS 223
179 SCDDAIKGVFNKFIHIGAVGIVAVVIMIFEMILSNVLCGIGINS 223
  SCDDAIKGVFNKFIHIGAVGIVAVVIMIFEMILSNVLCGIGINS 223

```

```

; Sequence 1, Application US/0840822B
; Patent No. 5776727
GENERAL INFORMATION:
APPLICANT: Ikeyama, Shuichi
APPLICANT: Koyama, Masaru
APPLICANT: Miyake, Masayuki
APPLICANT: Senoo, Masaharu
TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: Dike, Bronstein, Roberts & Cushman
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,222B
FILING DATE: 22-MAR-1995
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/254,493
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-079996-1991
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-085396-1991
FILING DATE: 14-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-022321-1992
FILING DATE: 07-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 41777-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 228 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal1
ORIGINAL SOURCE:
US-08-408-222B-1
Query Match 43.9%; Score 518.5; DB 1; Length 228;
Best Local Similarity 44.9%; Pred. No. 6,6e-52;
Matches 101; Conservative 39; Mismatches 72; Indels 13; Gaps 5;
QY 1 KYLLFVNFEMVLGGAGVILGVALMLRLDPOTTSILYLELGNKPAKPNTFYGIYLLIANGA 60
Db ||||| :||||| :|:||| :||| :|: ||||| :||| :||| :||| :||| :|||
QY 11 KYLLFGNPIFLMGLIANVAIGLMIRDSQTKSIFROETWNN--NSSFTTGVVILIGAGA 68
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 61 VMHFVGGLCYGAIOSSQCILGTFPTCLVLIPACEVAAGIWGFVNKKDIANKDYKOAYDQA 120
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 69 LMMLVGEPLGCCGAVQSQCMLGFEGFLVIIPAIRIAAAIMWSHKDEVLAKVGBPYKDT 128
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 121 LGQAVDDDNANNAKVVKTHRTLNCCGSAAALTTLTTTIUR--NSLCPSGNIITPLIQ 178
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 129 YNDATKTRDEQ--RETLKAIHVALNCG-----LAGVQKFIISDCPK-KDVLETFTVK 179
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

179 DCHOKIDELFSGLYLGIAIVAVAVIMIFEMILSMVLCGIRNS 223
180 SCPDAIKETFDNKHFIIGAVGIGIAVAVIMIFEMILSMVLCGIRN 224

LT 7

8-808-148-3
quence 3, Application US/08808148
tent No. 6020478

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer
APPLICANT: Goll, Surya
APPLICANT: Zhang, Hong Wolfe
TITLE OF INVENTION: NOVEL HUMAN TUMOR-ASSOCIATED ANTIGEN
NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/808.148

FILING DATE: Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0218 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 237 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 180926

1-808-148-3

1-808-148-3

Query Match 25.3%; Score 298; DB 3; Length 237;

Best Local Similarity 30.7%; Pred. No. 2.3e-26;

Matches 75; Conservative 45; Mismatches 86; Indels 38; Gaps 5;

1 KYLLFVNFVFMAGVYILGVALMLRHPDPTTSLYLIELGNKPAKPTFYVGIYIILVAVGA 60
9 KYSMPFNFPLFMCGIILALAIWVVSNDSDAI--FCSEBDVGSSTYAVVDILVAVGA 64
61 VMHVGGLGCGYAIOSOCILGTFPTCLVILFACVAVAGIWGFPVNDQIAKDKQ--FYD 118
65 IIMTLGLGCGGAIKESRCMLLPFGLILLILLQVAVYGLIAGVFKSKSRIVMETLYEN 124
119 QALQAVMODDANNAKAVVKTFFHETLNCSSNALTTTLTILNSLCPSGGNIL----- 172
125 TRLSANGSESEKQFOEHLI-VFOSEFCCG-----LVNGAADWGNPNFOHPELC 172
173 -----TPLLQODCHOKIDELFSGLYLGIAIVAVAVIMIFEMILSMVLCG 218
173 ACIDKORPCOSYNGKQVYKETSIFIKDFLAKNLIIVIGISFGIAVEIIGLVFSMWLYC 232

QY 219 GIRN 222
DB 233 QIGN 236

RESULT 8

US-08-855-140-4

Sequence 4, Application US/08855140

Patent No. 5854022

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: NEW TRANSMEMBRANE PROTEIN
NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/855.140

FILING DATE: Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0296 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 219 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1279546

US-08-855-140-4

Query Match 25.2%; Score 297.5; DB 2; Length 219;

Best Local Similarity 33.0%; Pred. No. 2.3e-26;

Matches 76; Conservative 38; Mismatches 85; Indels 31; Gaps 6;

1 KYLLFVNFVFMAGVYILGVALMLRHPDPTT-----SLYLIELGNKPAKPTFYVGIYI 54
10 KYLLFVNFPLFMCGCIIIGFYFLV-VQNTYGVLFRLMIFLITGN-----I 55
55 LIAVGAVMHVGGLGCGYAIOSOCILGTFPTCLVILFACVAVAGIWGFPVNDQIAKDKY 114
56 LVIVSGIINWVAFICGMSIKENKCLMSFVLLITILAEVYIALILFVYQKNTLVLA 115
115 QFYDQALQAVMODDANNAKAVVKTFFHETLNCSSNALTTTLTILNSLCPSGGNILTP 174
116 EGLNDSIQHYHSDNSTMKAMDPIQT---QLQCGVNGSSDWTSG--PSSCSGADV--- 167
QY 175 LLOQDCHOKIDELFSGLYLGIAIVAVAVIMIFEMILSMVLCGIRNS 224
DB 168 ---QCYNFAKSMFHSNPLFYIGITICVCVIOVLGMSFALTNCOIDKTS 214

1 9
 -430-225A-20
 jence 20, Application US/08430225A
 ent No. 6204000
 GENERAL INFORMATION:
 APPLICANT: Dong, Jin-Tang; Barrett,
 APPLICANT: J. Carl; Lamb, Patricia W.; Isaacs, John T.
 TITLE OF INVENTION: DIAGNOSTIC METHODS AND
 TITLE OF INVENTION: GENE THERAPY USING REAGENTS DERIVED FROM THE
 TITLE OF INVENTION: HUMAN METASTASIS SUPPRESSOR GENE KAI1
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNIGAN, L.L.P.
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/430,225A
 FILING DATE: 28-APR-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: RICHARD W. BORK
 REGISTRATION NUMBER: 36,459
 REFERENCE/DOCKET NUMBER: -2026-4172
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 TELEX: 421792
 FORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 267 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 -430-225A-20
 Y Match 23.1%; Score 272.5; DB 3; Length 267;
 : Local Similarity 26.3%; Pred. No. 2.4e-23;
 hes 68; Conservative 50; Mismatches 84; Indels 57; Gaps 7;
 1 KYLFVFNFFVFWLAGVILGVALLMRHDPQT-TSLLYLKGKPKAPNTFYVGIYILI 59
 10 KYFLFVFNLFVFLIGAVIILGFGVWILADKSSPISVL-----QTSSSLRMGAYVPIGVG 63
 60 AVMMFVFGICVCAIOESQCLIGFTFCVILFACGVAAIGVFNKQDIADVKQF 119
 64 AVTLMGFLGICCAVMBVRCLGLVFAFLLIILIAQVTAGALFFYFMGGLKQEMGI 123
 120 ALQ--QAVVDDDNANNAKAVVKTFFETLNCG----- 148
 124 LRDVNSSEDSIQDMDVYQ---QVKCGWVSFTWMDNMLMRPREVTPCSEVYG 180
 149 --SNALTTTTLTILNSLCPSGNLT-----DLQODCHQKIDELFSGLYILGIA 198
 181 EEDNSLS-----VRKGCEAPENRTQSGNHPEDMDVYQEGCKEKVQWILQENTLGIILGV 234
 199 AIVAVIIMIFEMILSVLC 217
 235 GVGVALIELKAVLSTICLC 253

; Patent No. 5854022
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Goll, Surya K.
 ; APPLICANT: Guegler, Karl J.
 ; TITLE OF INVENTION: NEW TRANSMEMBRANE PROTEIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/855,140
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0296 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 219 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 180141
 ; US-08-855-140-3
 Query Match 22.5%; Score 265.5; DB 2; Length 219;
 : Best Local Similarity 30.7%; Pred. No. 1.2e-22;
 Matches 70; Conservative 38; Mismatches 93; Indels 27; Gaps 5;
 QY 1 KYLFVFNFFVFWLAGVILGVALLMRHDPQTSLY-----LELGNKPKAPNTFYVGIYILI 56
 DB 10 KYFLFVFNLFVFLIGAVIILGFGVWILADKSSPISVL-----VFG 57
 QY 57 AVGAVMMFVFGICVCAIOESQCLIGFTFCVILFACGVAAIGVFNKQDIADVKQF 116
 DB 58 IGVSIIMVAVFLGCMGSIKENKCLMSFFILLIILAEVTAILLPYEQGLAEYVAKG 117
 QY 117 YDQALQAVVMDDNANNAKAVVKTFFETLNCGSNALTTTLTILNSLCPSGNLTPL 176
 DB 118 LTDSIHRYSN-----NSTRAANDSIOSFLQCGINGTSMTSG--PPACSPDRKV----- 167
 QY 117 OODCHQKIDELFSGLYILGIAIYVAVIIMIFEMILSVLCGIRNS 224
 DB 168 -EGYAKRLMFSNFLTITGIIITTCVIEVAGMSFALLNCOIDKTS 214
 RESULT 11
 US-08-807-044-3
 ; Sequence 3, Application US/08807044
 ; Patent No. 5863735
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.

ERAL INFORMATION:

PLICANT: Testa, Jacqueline B.
PLICANT: Quigley, James P.

PLICANT: Seandel, Marco

TITLE OF INVENTION: MONOCLONAL ANTIBODIES THAT RECOGNIZE ANTIGENS
TITLE OF INVENTION: ASSOCIATED WITH TUMOR METASTASIS

LE REFERENCE: SUNY

RENT APPLICATION NUMBER: US/09/333,599

RENT FILING DATE: 1999-06-15

MBER OF SEQ ID NOS: 5

FTWARS: Patent In Ver. 2.0

ID NO 4

ENGTH: 253

YPE: PRT

GANISM: Mus musculus

-333-599-4

Match 21.1%; Score 248.5; DB 3; Length 253;
Local Similarity 26.4%; Pred. No. 1.3e-20;
ches 65; Conservative 42; Mismatches 100; Indels 39; Gaps 7;

1 KYLFVFNFMVLAGVILGVALM--LRHDPQTSLLYLEGNKRPAPTFYVGIYILA 57

17 KYLFVFNFMVLAGVILGVALM--LRHDPQTSLLYLEGNKRPAPTFYVGIYILA 57

58 VGVNMFVFGFLGCGAIOESQCLLGFPTCLVILPACFVAGIWFVNDQIAKDYKQFY 117

67 AGVAVVAVTVLGGCCTAFKERRNLRLYFILLIIFLLEIAGIILAVVYQQLNTEIKENL 126

118 DQALQAVMDDDANNAKAVVKTFFHETLCCGSN-----ALTTLTILNSLCP 166

127 KQTVWVRVHQSHEGVSSAVDQLQDFHCSSNNSQDMSWIRSGEADSRVVPDSCCK 186

167 S-----GGNITPLLODCHQKIDELFSGLYLIGIAIYVAVIMIPEMILSMVL 216

187 TWVAGCGKRDHASNLYK--VEGGCTTKLETPIQEHRLVYIANGIGIACVQVGMIFT--- 241

217 CCGIN 222

242 CCLYRS 247

14

499-781-4

ence 4, Application US/09499781

ent No. 6498014

RAL INFORMATION:

PLICANT: Testa, Jacqueline B.

PLICANT: Quigley, James P.

PLICANT: Seandel, Marco

TITLE OF INVENTION: MONOCLONAL ANTIBODIES THAT RECOGNIZE ANTIGENS

TITLE OF INVENTION: ASSOCIATED WITH TUMOR METASTASIS

LE REFERENCE: SUNY

RENT APPLICATION NUMBER: US/09/499,781

RENT FILING DATE: 2000-02-08

OR APPLICATION NUMBER: 09/333,599

OR FILING DATE: 1999-06-15

BER OF SEQ ID NOS: 5

FTWARS: Patent In Ver. 2.0

ID NO 4

NGTH: 253

PR: PRT

GANISM: Mus musculus

499-781-4

Match 21.1%; Score 248.5; DB 4; Length 253;
Local Similarity 26.4%; Pred. No. 1.3e-20;
ches 65; Conservative 42; Mismatches 100; Indels 39; Gaps 7;

1 KYLFVFNFMVLAGVILGVALM--LRHDPQTSLLYLEGNKRPAPTFYVGIYILA 57

17 KYLFVFNFMVLAGVILGVALM--LRHDPQTSLLYLEGNKRPAPTFYVGIYILA 57

QY 58 VGVNMFVFGFLGCGAIOESQCLLGFPTCLVILPACFVAGIWFVNDQIAKDYKQFY 117

DB 67 AGVAVVAVTVLGGCCTAFKERRNLRLYFILLIIFLLEIAGIILAVVYQQLNTEIKENL 126

QY 118 DQALQAVMDDDANNAKAVVKTFFHETLCCGSN-----ALTTLTILNSLCP 166

DB 127 KQTVWVRVHQSHEGVSSAVDQLQDFHCSSNNSQDMSWIRSGEADSRVVPDSCCK 186

QY 167 S-----GGNITPLLODCHQKIDELFSGLYLIGIAIYVAVIMIPEMILSMVL 216

DB 187 TWVAGCGKRDHASNLYK--VEGGCTTKLETPIQEHRLVYIANGIGIACVQVGMIFT--- 241

QY 217 CCGIN 222

DB 242 CCLYRS 247

RESULT 15

US-08-705-771-19

Sequence 19, Application US/08705771

Patent No. 6054289

GENERAL INFORMATION:

APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,

TITLE OF INVENTION: Human Genes, Sequences and

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSER: CARELLA, BYRNE, BAIN, GILFILLAN,

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION NUMBER: US/08/705,771

FILING DATE: August 30, 1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

TELEPHONE: 973-994-1700

TELEFAX: 973-994-1744

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 236 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-705-771-19

Query Match 20.9%; Score 246.5; DB 3; Length 236;
Best Local Similarity 28.5%; Pred. No. 2e-20;
Matches 67; Conservative 44; Mismatches 105; Indels 19; Gaps 5;

QY 1 KYLFVFNFMVLAGVILGVALM--LRHDPQTSLLYLEGNKRPAPTFYVGIYILA 60

DB 10 KYLFVFNFMVLAGVILGVALM--LRHDPQTSLLYLEGNKRPAPTFYVGIYILA 62

QY 61 VGVNMFVFGFLGCGAIOESQCLLGFPTCLVILPACFVAGIWFVNDQIAKDYKQFYDQA 120

DB 63 FVMAIGFVGCIGAIKCKKCLLTFILLIIFLLEIAGIILAVVYQQLNTEIKENL 122

QY 121 LQAVVMDDDANNAKAVVKTFFHETLCCGSNALT---LTTTILNSLCP---PSCGNT 171

123 LHXGTQANGLTNANSIIOT---DFRCGVSNYTDWFEVYNATRVDPDSCCLEFSESQGL 179
172 LPELLOQDCHQKIDELFSGRLYLIGIAIVAVIMIPEMILSMVLCCGIRNSSVY 226
180 HAPALVEGRATRGEVAGSGLAVGIFGLCTALVOLIGLNFAMTMTWQVVKADTY 234

ch completed: December 9, 2003, 10:33:33
time : 31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

protein - protein search, using sw model

on: December 9, 2003, 10:32:07 ; Search time 350 Seconds
(without alignments)
120.092 Million cell updates/sec

ect score: US-10-035-914-2

ence: 1 KYLLFVNFVFLAGVILG.....IFEMILSMVLCGGIRNSSVY 226

ing table: BLOSUM62

ched: Gapop 10.0 , Gapext 0.5

1 number of hits satisfying chosen parameters: 684280

mm DB seq length: 0

mm DB seq length: 200000000

-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

base : Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PTCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Score	Query	Match	Length	ID	Description
1	1103	93.5	236	11	US-09-999-121-5 Sequence 5, Appli
2	1103	93.5	236	12	US-10-169-297-7 Sequence 7, Appli
3	1034	87.6	262	9	US-09-925-299-794 Sequence 794, App
4	1034	87.6	262	11	US-09-925-299-794 Sequence 794, App
5	1034	87.6	279	10	US-09-731-872-396 Sequence 396, App
6	1034	87.6	279	12	US-09-876-997-396 Sequence 396, App
7	522.5	44.3	226	12	US-10-205-194-162 Sequence 162, App
8	518.5	43.9	227	8	US-08-592-711-6 Sequence 6, Appli
9	518.5	43.9	227	11	US-09-350-202-6 Sequence 6, Appli
0	518.5	43.9	228	16	US-10-156-136-39 Sequence 39, Appli
1	518.5	43.9	275	9	US-09-925-301-1381 Sequence 1381, Ap
2	518.5	43.9	275	15	US-10-106-698-5930 Sequence 5930, Ap
3	462.5	39.2	221	12	US-10-823-219-185 Sequence 185, App
4	453.5	38.4	221	9	US-09-823-356-17 Sequence 17, Appli
5	453.5	38.4	221	16	US-10-156-136-25 Sequence 25, Appli

16	453.5	38.4	229	15	US-10-106-698-4519 Sequence 4519, Ap
17	400	33.9	209	15	US-10-106-698-6825 Sequence 6825, Ap
18	339	28.7	157	11	US-09-823-187-76 Sequence 76, Appli
19	316	26.8	254	10	US-09-934-268-4 Sequence 4, Appli
20	316	26.8	254	15	US-10-162-435-38 Sequence 38, Appli
21	298	25.3	237	9	US-09-923-217-1062 Sequence 1062, Ap
22	298	25.3	237	10	US-09-833-263-1062 Sequence 1062, Ap
23	298	25.3	237	14	US-10-025-380-1062 Sequence 1062, Ap
24	298	25.3	241	9	US-09-925-301-1035 Sequence 1035, Ap
25	298	25.3	241	15	US-10-106-698-5105 Sequence 5105, Ap
26	298	25.3	243	9	US-09-922-217-1122 Sequence 1122, Ap
27	298	25.3	243	14	US-10-025-380-1122 Sequence 1122, Ap
28	298	25.3	446	9	US-09-922-217-1121 Sequence 1121, Ap
29	298	25.3	446	14	US-10-025-380-1121 Sequence 1121, Ap
30	297.5	25.2	219	12	US-09-976-782-118 Sequence 118, App
31	289.5	24.5	249	10	US-09-908-193-16 Sequence 16, Appli
32	288	24.4	248	10	US-09-908-193-41 Sequence 41, Appli
33	288	24.4	248	15	US-10-103-196-15 Sequence 15, Appli
34	286.5	24.3	219	12	US-09-576-782-117 Sequence 117, App
35	286.5	24.3	219	16	US-10-156-136-34 Sequence 34, Appli
36	283.5	24.0	247	10	US-09-908-193-42 Sequence 42, Appli
37	283.5	24.0	247	11	US-09-823-187-72 Sequence 72, Appli
38	282.5	23.9	239	9	US-09-823-356-14 Sequence 14, Appli
39	272.5	23.1	239	12	US-09-876-782-119 Sequence 119, App
40	272.5	23.1	267	9	US-09-795-380-20 Sequence 20, Appli
41	267.5	22.7	267	11	US-09-823-187-73 Sequence 73, Appli
42	265.5	22.5	219	11	US-09-823-187-75 Sequence 75, Appli
43	265.5	22.5	219	12	US-09-970-424-3 Sequence 3, Appli
44	265.5	22.5	219	12	US-09-976-782-116 Sequence 116, Appli
45	265.5	22.5	231	9	US-09-925-302-474 Sequence 474, App

ALIGNMENTS

RESULT 1
US-09-999-121-5
Sequence 5, Application US/09999121
Publication No. US20030039982A1
GENERAL INFORMATION:
APPLICANT: Ryan, James W.
TITLE OF INVENTION: ISOLATED GENOMIC NUCLEOTIDE FRAGMENTS FROM THE p15 REGION OF
FILE REFERENCE: US-10,001-US
CURRENT APPLICATION NUMBER: US/09/999,121
CURRENT FILING DATE: 2002-05-29
PRIORITY APPLICATION NUMBER: US 60/244,705
PRIORITY FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 236
TYPE: PRT
ORGANISM: Homo sapiens
US-09-999-121-5

Query Match 93.5%; Score 1103; DB 11; Length 236;
Best Local Similarity 91.6%; Pred. No. 1.6e-107;

Matches 207; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

QY 1 KYLLFVNFVFLAGVILGVALMLRHPDPTSLYLRLGNKPAFTFYGIYILAVGA 60
DB 11 KYLLFVNFVFLAGVILGVALMLRHPDPTSLYLRLGNKPAFTFYGIYILAVGA 70
QY 61 VMVFVFLGCTGATGSOCLGTFCTVILFACFVAAGIWKGVNDDOIAKYKOTYDOA 120
DB 71 VMVFVFLGCTGATGSOCLGTFCTVILFACFVAAGIWKGVNDDOIAKYKOTYDOA 130
QY 121 LQAAVDDDDANNAKAAVKTFFHETLNCSSNALTTTLTILRNLSLCPSGNILLPLLAQDC 180
DB 131 LQAAVDDDDANNAKAAVKTFFHETLNCSSNALTTTLTILRNLSLCPSGNILLPLLAQDC 190
QY 181 HOKIDELFSGKLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 226

|||||
191 HOKIDDLFSGKLYLIGIAIVAVIMIFEMILSMVLCGGIRNSSVY 236

T 2
-169-297-7
ence 7, Application US/10169297
lication No. US20030171276A1
ERAL INFORMATION:

PPLICANT: Tohdoh, Naoki

PPLICANT: Murata, Masashi

PPLICANT: Enjoji, Takashi

FILE OF INVENTION: Preventives and Remedies for Chronic

FILE OF INVENTION: Hepatitis

LE REFERENCE: 3435.1000-000

RENT APPLICATION NUMBER: US/10/169,297

RENT FILING DATE: 2002-10-31

FOR APPLICATION NUMBER: PCT/JP00/09393

FOR FILING DATE: 2000-12-28

FOR APPLICATION NUMBER: JP 11/374087

FOR FILING DATE: 1999-12-28

MBER OF SEQ ID NOS: 50

FTWARE: FastSeq for Windows Version 4.0

ID NO 7

ENGTH: 236

YPE: PRT

GANISM: Homo sapiens

-169-297-7

Match 93.5%; Score 1103; DB 12; Length 236;
Local Similarity 91.6%; Pred. No. 1.6e-107;
ches 207; Conservative 14; Mismatches 5; Indels 0; Gaps 0;
1 KTLVFNVLAVAGVILGVALMLRHPDPTSLYLSEKRPAPTFYGIYLLAVGA 60
11 KTLVFNVLAVAGVILGVALMLRHPDPTSLYLSEKRPAPTFYGIYLLAVGA 70
61 VMMFVGLGCTGALIGSQCILGFFTCVILPACVAAAGIMGFVNDQIAKQKPYDQA 120
71 VMMFVGLGCTGALIGSQCILGFFTCVILPACVAAAGIMGFVNDQIAKQKPYDQA 130
121 LQAAVDDDDANNAKAVVKTFFHETLNCSSNALTTTLTILRNSLCPSGNITLPLLODC 180
131 LQAAVDDDDANNAKAVVKTFFHETLNCSSNALTTTLTILRNSLCPSGNITLPLLODC 190
181 HOKIDDLFSGKLYLIGIAIVAVIMIFEMILSMVLCGGIRNSSVY 226
191 HOKIDDLFSGKLYLIGIAIVAVIMIFEMILSMVLCGGIRNSSVY 236

F 3
-925-299-794
ence 794, Application US/09925299
ant No. US20020055627A1
ERAL INFORMATION:

PPLICANT: Rosen et al.

FILE OF INVENTION: Nucleic Acids, Proteins and Antibodies

RE REFERENCE: Pal02

RENT APPLICATION NUMBER: US/09/925,299

RENT FILING DATE: 2001-08-10

FOR APPLICATION NUMBER: PCT/US00/05883

FOR FILING DATE: 2000-03-08

FOR APPLICATION NUMBER: 60/124,270

FOR FILING DATE: 1999-03-12

MBER OF SEQ ID NOS: 1556

FTWARE: PatentIn Ver. 2.0

ID NO 794

ENGTH: 262

YPE: PRT

GANISM: Homo sapiens

ZATURE:

ME/KEY: SITE

CATION: (38)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-794

Query Match 87.6%; Score 1034; DB 9; Length 262;
Best Local Similarity 91.1%; Pred. No. 3.3e-100;
Matches 195; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

QY 13 LAGVILGVALMLRHPDPTSLYLSEKRPAPTFYGIYLLAVGAVMMFVGLGCTG 72
DB 49 LAGVILGVALMLRHPDPTSLYLSEKRPAPTFYGIYLLAVGAVMMFVGLGCTG 108
QY 73 AIOESQCILGFFTCVILPACVAAAGIMGFVNDQIAKQKPYDQALQAAVDDDDANN 132
DB 109 AIOESQCILGFFTCVILPACVAAAGIMGFVNDQIAKQKPYDQALQAAVDDDDANN 168
QY 133 AKAVKTFFHETLNCSSNALTTTLTILRNSLCPSGNITLPLLODCQKIDBLFSGKL 192
DB 169 AKAVKTFFHETLNCSSNALTTTLTILRNSLCPSGNITLPLLODCQKIDBLFSGKL 228
QY 193 YLIGIAIVAVIMIFEMILSMVLCGGIRNSSVY 226
DB 229 YLIGIAIVAVIMIFEMILSMVLCGGIRNSSVY 262

RESULT 4
US-09-925-299-794
Sequence 794, Application US/09925299
Publication No. US20030040617A9
GENERAL INFORMATION:
PPLICANT: Rosen et al.
FILE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: Pal02
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 794
LENGTH: 262
TYPE: PRT
GANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (38)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-794

Query Match 87.6%; Score 1034; DB 11; Length 262;
Best Local Similarity 91.1%; Pred. No. 3.3e-100;
Matches 195; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

QY 13 LAGVILGVALMLRHPDPTSLYLSEKRPAPTFYGIYLLAVGAVMMFVGLGCTG 72
DB 49 LAGVILGVALMLRHPDPTSLYLSEKRPAPTFYGIYLLAVGAVMMFVGLGCTG 108
QY 73 AIOESQCILGFFTCVILPACVAAAGIMGFVNDQIAKQKPYDQALQAAVDDDDANN 132
DB 109 AIOESQCILGFFTCVILPACVAAAGIMGFVNDQIAKQKPYDQALQAAVDDDDANN 168
QY 133 AKAVKTFFHETLNCSSNALTTTLTILRNSLCPSGNITLPLLODCQKIDBLFSGKL 192
DB 169 AKAVKTFFHETLNCSSNALTTTLTILRNSLCPSGNITLPLLODCQKIDBLFSGKL 228
QY 193 YLIGIAIVAVIMIFEMILSMVLCGGIRNSSVY 226
DB 229 YLIGIAIVAVIMIFEMILSMVLCGGIRNSSVY 262

RESULT 5
US-09-731-872-396

Sequence 396, Application US/09731872
 Patent No. US20020102604A1
 GENERAL INFORMATION:
 APPLICANT: Dumas Milne Edwards, Jean Baptiste
 APPLICANT: Bouquelerec, Lydie
 APPLICANT: Jobert, Severin
 TITLE OF INVENTION: FULL-LENGTH HUMAN CDNA5 ENCODING POTENTIALLY SECRETED PROTEINS
 FILE REFERENCE: 78.053.REG
 CURRENT APPLICATION NUMBER: US/09/731,872
 PRIOR FILING DATE: 2000-12-07
 PRIOR APPLICATION NUMBER: US 60/169,629
 PRIOR FILING DATE: 1999-12-08
 PRIOR APPLICATION NUMBER: US 60/187,470
 PRIOR FILING DATE: 2000-03-06
 OPTIMAR: Patent.pm
 Q ID NO 396
 LENGTH: 279
 TYPE: PRT
 ORGANISM: Homo sapiens
 9-731-872-396

Query Match 87.6%; Score 1034; DB 10; Length 279;
 at Local Similarity 91.1%; Pred. No. 3.5e-100;
 ches 195; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

13 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 72
 66 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 125
 73 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 132
 126 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 185
 133 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 192
 186 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 245
 193 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 226
 246 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 279

UT 6
 3-876-997-396
 Sequence 396, Application US/09876997
 Application No. US20030152821A1
 GENERAL INFORMATION:
 APPLICANT: Dumas Milne Edwards, Jean Baptiste
 APPLICANT: Bouquelerec, Lydie
 APPLICANT: Jobert, Severin
 TITLE OF INVENTION: FULL-LENGTH HUMAN CDNA5 ENCODING POTENTIALLY SECRETED PROTEINS
 FILE REFERENCE: 78.054.CIP
 CURRENT APPLICATION NUMBER: US/09/876,997
 PRIOR FILING DATE: 2001-06-08
 PRIOR APPLICATION NUMBER: US 09/731,872
 PRIOR FILING DATE: 2000-12-07
 PRIOR APPLICATION NUMBER: US 60/187,470
 PRIOR FILING DATE: 2000-03-06
 PRIOR APPLICATION NUMBER: US 60/169,629
 PRIOR FILING DATE: 1999-12-08
 OPTIMAR: Patent.pm
 Q ID NO 396
 LENGTH: 279
 TYPE: PRT
 ORGANISM: Homo sapiens
 3-876-997-396

Query Match 87.6%; Score 1034; DB 12; Length 279;
 at Local Similarity 91.1%; Pred. No. 3.5e-100;
 ches 195; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

Query
 13 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 72
 66 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 125
 73 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 132
 126 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 185
 133 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 192
 186 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 245
 193 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 226
 246 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 279

Db
 13 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 72
 66 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 125
 73 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 132
 126 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 185
 133 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 192
 186 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 245
 193 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 226
 246 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 279

Query
 13 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 72
 66 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 125
 73 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 132
 126 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 185
 133 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 192
 186 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 245
 193 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 226
 246 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 279

Db
 13 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 72
 66 LAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGAVMMFVGLGCG 125
 73 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 132
 126 AIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQAAMDNDANN 185
 133 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 192
 186 AKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLODCHQKIDELFSGKL 245
 193 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 226
 246 YLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 279

Query Match 44.3%; Score 522.5; DB 12; Length 226;
 Best Local Similarity 45.1%; Pred. No. 1.2e-46;
 Matches 101; Conservative 39; Mismatches 71; Indels 13; Gaps 5;

1 KYLIFVNFVFMLAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGA 60
 11 KYLIFVNFVFMLAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGA 66
 61 VMMFVGLGCGAIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQA 120
 67 LMMVGLGCGAIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQA 126
 121 LQQAAMDNDANNNAKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLOD 179
 127 YOKLRNDDEPO--RETUKAITHMALMCCGIAGVGEQFISDI-----CPK-KQVLESFQVKS 178
 180 CHQKIDELFSGKLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 223
 179 CPDAIDVFNHSHKHITIGAVGIGIAVVMIFEMIFSMILCAIRRS 222

Db
 1 KYLIFVNFVFMLAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGA 60
 11 KYLIFVNFVFMLAGVILGVALMLRHPDPTSLIYLELGNKRPAPNTFYVGIYILINAVGA 66
 61 VMMFVGLGCGAIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQA 120
 67 LMMVGLGCGAIOESQCLGTFETCLVILIFACEVAAGIMGFVFNKQIADVQKQFDQALQQA 126
 121 LQQAAMDNDANNNAKAVVKTETHTLNCSSNALITLTTILNLSLCPGSGNITLPLLOD 179
 127 YOKLRNDDEPO--RETUKAITHMALMCCGIAGVGEQFISDI-----CPK-KQVLESFQVKS 178
 180 CHQKIDELFSGKLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 223
 179 CPDAIDVFNHSHKHITIGAVGIGIAVVMIFEMIFSMILCAIRRS 222

RESULT 8
 US-08-592-711-6
 Sequence 6, Application US/08592711
 Publication No. US20020115214A1
 GENERAL INFORMATION:
 APPLICANT: June, Carl H.
 APPLICANT: Thompson, Craig B.
 APPLICANT: Nabel, Gary J.
 APPLICANT: Gray, Gary S.

APPLICANT: Remert, Paul D.
TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVA & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,711
FILING DATE: 26-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/435,816
FILING DATE: 4-MAY-1995
APPLICATION NUMBER: US 08/403,253
FILING DATE: 10-MARCH-1995
APPLICATION NUMBER: US 08/253,964
FILING DATE: 3-JUNE-1994
APPLICATION NUMBER: US 08/073,223
FILING DATE: 4-JUNE-1993
APPLICATION NUMBER: US 08/200,947
FILING DATE: 23-FEB-1994
APPLICATION NUMBER: US 07/864,805
FILING DATE: 7-APR-1992
APPLICATION NUMBER: US 08/247,505
FILING DATE: 23-MAY-1994
APPLICATION NUMBER: US 07/864,866
FILING DATE: 7-APR-1992
APPLICATION NUMBER: US 08/218,155
FILING DATE: 25-MAR-1994
APPLICATION NUMBER: US 07/864,807
FILING DATE: 7-APR-1992
APPLICATION NUMBER: US 07/902,467
FILING DATE: 16-JUNE-1992
APPLICATION NUMBER: US 07/275,433
FILING DATE: 23-NOV-1988

TORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy B.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPI-002CP4
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
FORMATION FOR SBO ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
592-711-6

Y Match 43.9%; Score 518.5; DB 8; Length 227;
Local Similarity 44.9%; Pred. No. 3,2e+46;
Res 101; Conservative 39; Mismatches 72; Indels 13; Gaps 5;

1 KYLLPFNPNFMILAGVVIIQVALMLRHDDQTSLYLIELGNPDAPNTFYGVIIYLIAVGA 60
||| ||| : | : | : | : | : | : | : | : | : | : | : | : |
10 KYLLPFGNFIFMLGAVIALAIGMLRFDSQTSIPEQETNNN--NSSPTTGYIILIGGA 67
::| ||| ||| : | : | : | : | : | : | : | : | : | : | : | : |
68 LMLVGFGLGCCAVOESQCMLGISFGFLIYIPALIRIAAIWGSYKSDIVIKVOEFYMDT 127
::| ||| ||| : | : | : | : | : | : | : | : | : | : | : | : |
121 LOOAWMDDANNNAKAAVKTFPHETLINCSSNALTTTLITTLIR--NSLCPSGGNIITPLIQ 178
||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
128 YNLKTKRDDEQ--RETLAKAHYALNCCG-----LAGVEQISIDICPK-KDVLSTFTVK 178

```

09 179 DOHOKIDELFSGKLYLIGIAIVAVIMFEMILSWLCCGINS 223
| | | | | : | | | | | : | | | | | : | | | | | :
DB 179 SCPDAKEVFDMKFFHILIGAVGIGIAVIMFGMIFSMILCAIRN 223

RESULT 9
US-09-350-202-6
; Sequence 6, Application US/09350202
; Publication No. US2003009643A1
GENERAL INFORMATION:
APPLICANT: June, Carl H., Thompson, Craig B., Nabel, Gary J.
APPLICANT: Gray, Gary S., Rennett, Paul D.
TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/350,202
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/403,253
FILING DATE: March 10, 1995
APPLICATION NUMBER: US 08/253,964
FILING DATE: 3 JUNE 1994
APPLICATION NUMBER: US 08/073,223
FILING DATE: 4 JUNE 1993
APPLICATION NUMBER: US 08/200,947
FILING DATE: 23 FEB 1994
APPLICATION NUMBER: US 07/864,805
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 08/247,505
FILING DATE: 23 MAY 1994
APPLICATION NUMBER: US 07/864,866
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 08/218,155
FILING DATE: 25 MAR 1994
APPLICATION NUMBER: US 07/864,807
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 07/902,467
FILING DATE: 16 JUNE 1992
APPLICATION NUMBER: US 07/275,433
FILING DATE: 23 NOV 1988
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPI-002CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SRO ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-350-202-6

```


OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/156,136
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/010,146
FILING DATE: <Unknown>
APPLICATION NUMBER: WO US98/00959
FILING DATE: 21-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF354PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
FAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
0-156-136-25

Seq Match 38.4%, Score 453.5, DB 16, Length 221;
Local Similarity 41.3%, Pred. No. 2e-39;
tches 92; Conservative 33; Mismatches 81; Indels 17; Gaps 4;

```
1 KILFVFNFMAGVILGVVALMLRHPQTSLYLBIANKPAPRTFVGYIILAVCA 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
12 KILLIGFNLFLASAVAFGLMFRFGAIKEL-----SSDKSPRYFVGLYLVAGGA 67
61 VMFVGLGCGYCAIESQCLGTFFCTVILFACVAAAGVGVNNDQIAKDVQFYDOA 120
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 LMAVGFPGCCGAMESQCVLSFTCLVIFAEVTVGVFAFVGKVAIRHVQMTYEA 127
121 LQQAIVDDANNAKAVKTFHETLNCSSNALTTTLIRNSLCPSGGNLTPLLODC 180
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 YNDYKDRGKNGTLI--TFHSTFOCCGKSSSEGVQPT-----CPK-----EILGHKNC 174
181 HOKIDELPSGKLYLGIAIYAVAVMIFEMILSMTLCCGIRNS 223
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175 IDRIETISVRLQILGIVGIGIAGLTFGMIFSMVLCALIRNS 217
```

Job completed: December 9, 2003, 11:10:20
Time : 352 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comogen Ltd.

rotein - protein search, using sw model

on: December 9, 2003, 09:56:34 / Search time 36 Seconds
(without alignments)

603.726 Million cell updates/sec

ect score: US-10-035-914-2

ence: 1 KYLLFVFNFMVLAGVILG.....IFEMILSMVLCGGIRNSSVY 226

ing table: BLOSUM62
Gapop 10.0 , Gapext 0.5

ched: 283308 segs, 96168682 residues

1 number of hits satisfying chosen parameters: 283308

num DB seq length: 0
num DB seq length: 2000000000

-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

ase : PIR_76:.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	ID	Description
1 1177	99.7	236	A46472	cell surface prote
2 1103	93.5	236	A35649	cell surface prote
3 528.5	44.8	226	UX0221	CD9 antigen - bovi
4 522.5	44.3	226	S39262	CD9 antigen - rat
5 518.5	43.9	228	A40402	CD9 antigen [valid
6 512.5	43.4	228	A42929	CD9 antigen - gree
7 510.5	43.3	226	I49589	antigen - mouse
8 420	35.6	222	A59253	tetraspan TSPAN-2
9 298	25.3	237	A36056	tumor-associated a
0 286.5	24.3	219	A39574	leukocyte antigen
1 272.5	23.1	267	A46493	metastasis suppres
2 265.5	22.5	219	A37243	hemopoietic cell s
3 262	22.2	266	I49561	C33/R2/TA4 - mouse
4 255.5	21.7	282	T21596	hypothetical prote
5 254.5	21.6	228	A59265	tetraspan TSPAN-4
6 227.5	19.3	233	T15620	hypothetical prote
7 221.5	18.8	241	A59262	tetraspan TSPAN-1
8 218.5	18.5	223	T26763	hypothetical prote
9 209	17.7	244	I39368	T-cell acute lymph
0 207	17.5	238	I43511	CD63/MR91 antigen
1 205	17.4	245	A59258	tetraspan TSPAN-6
2 199.5	16.9	254	T13615	hypothetical prote
3 196	16.6	253	A59264	tetraspan TSPAN-3
4 195	16.5	245	A59260	tetraspan TSPAN-6
5 194.5	16.5	248	A46508	CD63/MR91 antigen
6 191	16.2	281	A47629	cell surface glyco
7 187	15.8	238	A59261	tetraspan TSPAN-5
8 180.5	15.3	281	UC2297	CD63 antigen - rab
9 180.5	15.3	281	B47629	cell surface glyco

30	178.5	15.1	308	2	T24912	hypothetical prote
31	178	15.1	242	2	T15361	hypothetical prote
32	174.5	14.8	206	2	T25161	hypothetical prote
33	170	14.4	218	1	A43522	23k integral membr
34	170	14.4	238	1	I38016	melanoma - associa
35	162	13.7	321	2	T45053	hypothetical prote
36	158	13.4	218	1	A40181	23k integral membr
37	149.5	12.7	194	2	T25548	hypothetical prote
38	141.5	12.0	394	2	T18752	hypothetical prote
39	113	9.6	203	2	T22537	hypothetical prote
40	113	9.6	263	2	T02906	senescence-associa
41	109	9.2	427	2	T32652	hypothetical prote
42	106	9.0	244	2	S44610	C02F5.8 protein -
43	105	8.9	260	2	T46081	uroplakin Ib - bov
44	100.5	8.5	359	2	T18667	hypothetical prote
45	100	8.5	258	2	I46080	uroplakin Ia - bov

ALIGNMENTS

RESULT 1
A46472
cell surface protein TAPA-1 - mouse
N/Alternate names: target of antiProliferative antibody (TAPA) 1
C/Species: Mus musculus (house mouse)
C/Date: 18-Jun-1993 #sequence_revision 09-Aug-1996 #text_change 22-Jun-1999
C/Accession: A46472
R/Andrta, M.L.; Hsieh, C.L.; Oren, R.; Francke, U.; Levy, S.
J. Immunol. 147, 1030-1036, 1991
A/Title: Genomic organization and chromosomal localization of the TAPA-1 gene.
A/Reference number: A46472; MIMD:91318144; PMID:1650385
A/Molecule type: DNA
A/Residues: 1-236 <AND>
A/Cross-references: GB:845012; NID:91679982; PIDN:RA819417.1; PID:9233253
A/Experimental source: B-cell lymphoma line 38C13
A/Note: Sequence extracted from NCBI backbone (NCBIN:44957, NCBIN:44966, NCBIN:45001, NCBI
C/Superfamily: CD9 antigen
C/Keywords: transmembrane protein
F/2-11/Domain: intracellular #status predicted <CY1>
F/12-35/Domain: transmembrane #status predicted <TM1>
F/36-57/Domain: extracellular #status predicted <EX1>
F/58-84/Domain: transmembrane #status predicted <TM2>
F/85-113/Domain: intracellular #status predicted <CY2>
F/89-113/Domain: transmembrane #status predicted <TM3>
F/114-203/Domain: extracellular #status predicted <EX2>
F/204-230/Domain: transmembrane #status predicted <TM4>
F/231-236/Domain: intracellular #status predicted <CY3>

Query Match 99.7%; Score 1177; DB 1; Length 236;
Best Local Similarity 99.6%; Pred. No. 1e-97;

Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	KYLLFVFNFMVLAGVILGVALMLRHPDQTTSLYLBLEGNRPAPTFVVGIIYLLAVGA	60
DB	11	KYLLFVFNFMVLAGVILGVALMLRHPDQTTSLYLBLEGNRPAPTFVVGIIYLLAVGA	70
QY	61	VMMFVGFLGCGAIGSOCLLGTFFPCVLVILFACRFAAGINGFANQDQIAKYVKQFYDQA	120
DB	71	VMMFVGFLGCGAIGSOCLLGTFFPCVLVILFACRFAAGINGFANQDQIAKYVKQFYDQA	130
QY	121	LQAAVMDDDANNAKAVKTFHETLNCSSNALTTTLTLRLSLCPGSGNITLPLLOQDC	180
DB	131	LQAAVMDDDANNAKAVKTFHETLNCSSNALTTTLTLRLSLCPGSGNITLPLLOQDC	190
QY	181	HOKIDELPSGKLYLIGIAIVAVIMIFEMILSMVLCGGIRNSSVY	226
DB	191	HOKIDELPSGKLYLIGIAIVAVIMIFEMILSMVLCGGIRNSSVY	236
RESULT 2		A35649	

LT 5

antigen [validated] - human
ternate names: motility-related protein-1
esies: Homo sapiens (man)
te: 06-Dec-1991 #sequence, revision 07-Jul-1995 #ext_change 08-Dec-2000
cession: A46123; A40402; JH0555; A39029; S10564
miner 16, 132-138, 1993
cle: Organization of the human CD9 gene.
ference number: A46123; MUID:93252369; PMID:8486348
cession: A46123
lecular type: DNA
idues: 1-228 <RMB>
ss-references: GB:IJ34068; NID:g300112; PIDN:AAC0586.1; PID:g300115
erimental source: leukocyte
ze: sequence extracted from NCBI backbone (NCBIN:131318, NCBIN:131326, NCBIN:131328,
za, P.; Wolf, D.; Fox, C.F.; Kieffer, N.; Seyer, J.M.; Fried, V.A.; Coughlin, S.R.,
ol. Chem. 266, 10658-10645, 1991
le: cDNA cloning and expression of platelet p24/CD9. Evidence for a new family of r
ference number: A40402; MUID:91244846; PMID:2037603
ession: A40402
tus: not compared with conceptual translation
lecular type: mRNA
idues: 1-228 <LMAN>
ss-references: GB:IJ34068; GB:M61880; NID:g508495; PIDN:AAA5982.1; PID:g508496
e: parts of this sequence, including the amino end of the mature protein, were conf
ake, M.; Koyama, M.; Seno, M.; Ikegawa, S.
p. Med. 174, 1347-1354, 1991
le: Identification of the motility-related protein (MRP-1), recognized by monoclonal
reference number: JH0555; MUID:92078843; PMID:1720807
ession: JH0555
ecule type: mRNA
idues: 1-228 <MTY>
ss-references: GB:X60111; NID:g34768; PIDN:CMA2708.1; PID:g34769
erimental source: breast carcinoma
e: this protein has the epitope defined by cell motility-inhibiting monoclonal anti
chickx, C.; Benoit, P.; Prachet, P.; Billard, M.; Worthington, R.E.; Gagnon, J.; Uga
ol. Chem. 266, 117-122, 1991
le: Molecular cloning of the CD9 antigen. A new family of cell surface proteins.
reference number: A39029; MUID:91093112; PMID:1840589
ession: A39029
ecule type: mRNA
idues: 1-8, 'S', 10-66, 'A', 68-193, 195-228 <BOU>
ss-references: GB:M38690
e: parts of this sequence, including the amino end of the mature protein, were conf
ashahar, M.; Takahara, K.; Yacomi, Y.; Nakahara, K.; Kurokawa, K.
lett. 264, 270-274, 1990
le: Purification and partial characterization of CD9 antigen of human platelets.
reference number: S10564; MUID:90292223; PMID:2358073
ession: S10564
ecule type: protein
idues: 2-8, 'X', 10-21 <HTG>
etics:
e: GDB:CD9; MITC3
ss-references: GDB:120582; OMIM:143030
osition: 12p13-12p13
xtfamily: CD9 antigen
words: glycoprotein; transmembrane protein
28/Product: CD9 antigen #status experimental <MAT>
1/Domain: intracellular #status predicted <CY1>
35/Domain: membrane #status predicted <TM1>
55/Domain: extracellular #status predicted <EX1>
82/Domain: transmembrane #status predicted <TM2>
86/Domain: intracellular #status predicted <CY2>
111/Domain: transmembrane #status predicted <TM3>
194/Domain: extracellular #status predicted <EX2>

Query Match 43.9%, Score 518.5, DB 1, Length 228, Best Local Similarity 44.9%, Pred. No. 6, 1e-39, Matches 101, Conservative 39, Mismatches 72, Indels 13, Gaps 5;

QY 1 KYLLFVFNFMVLAAGVILGVALMLRHPDQTSLLYLELGNKRPAPTFYVGIYILAVGA 60
DB 11 KYLLFVFNFMVLAAGVILGVALMLRHPDQTSLLYLELGNKRPAPTFYVGIYILAVGA 68
QY 61 VMHVFVGFICGCAIQBSOCLGTFITFCVLIIIFACEVAAGIWGFVNKDQIAKDVQPYDQA 120
DB 69 LMMVLVGFICGCAIQBSOCLGTFITFCVLIIIFACEVAAGIWGFVNKDQIAKDVQPYDQA 128
QY 121 LQGVAMDDANNAKAVVKTFFHETLNCSSNALTTTLTLTLR--NSLCPSGNITLPLQQ 178
DB 129 YNKLKTRDPR--RETLLKAIHYALNCCG-----LAGVRFQSLSDICPK-KDVLFTFTVK 179
QY 179 DGHQKIDELFSGKLYLIGIAIVAVVIMIFEMILSMVLCGGIRNS 223
DB 180 SCPDAIKVEFDNKKFIIIGAVGIGIAVVMIFGMIFSMILCCAIRN 224

RESULT 6
A42929
CD9 antigen - green monkey
N/Alternate names: 27K diptheria toxin receptor-associated protein DRAP27
C/Species: Cercopithecus aethiops (green monkey, grivet)
C/Date: 01-Oct-1992 #sequence_revision 09-Aug-1996 #text_change 16-Jun-2000
C/Accession: A42929
R/Mitamura, T.; Yamoto, R.; Umata, T.; Yomo, T.; Urabe, I.; Tsunoeoka, M.; Mekada, E.
J. Cell Biol. 118: 1189-1199, 1992
A/Title: The 27-kD diptheria toxin receptor-associated protein (DRAP27) from vero cells
epitopes on toxin-sensitive cells.
A/Reference number: A42929; MUID:92394967; PMID:1522113
A/Accession: A42929
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-228 <MT>
A/Cross-references: GB:D10726; NID:g218565; PIDN:BAA01569.1; PID:g218566
C/Superfamily: CD9 antigen
C/Keywords: glycoprotein; transmembrane protein
F:2-328/Product: CD9 antigen #status predicted <MAT>
F:2-11/Domain: intracellular #status predicted <CY1>
F:12-35/Domain: transmembrane #status predicted <TM1>
F:36-55/Domain: extracellular #status predicted <EX1>
F:56-82/Domain: transmembrane #status predicted <TM2>
F:83-86/Domain: intracellular #status predicted <CY2>
F:87-111/Domain: transmembrane #status predicted <TM3>
F:112-194/Domain: extracellular #status predicted <EX2>
F:195-221/Domain: transmembrane #status predicted <TM4>
F:222-228/Domain: intracellular #status predicted <CY3>
F:52/53/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 43.4%, Score 512.5, DB 1, Length 228, Best Local Similarity 44.4%, Pred. No. 2, 1e-38, Matches 100, Conservative 40, Mismatches 72, Indels 13, Gaps 5;

QY 1 KYLLFVFNFMVLAAGVILGVALMLRHPDQTSLLYLELGNKRPAPTFYVGIYILAVGA 60
DB 11 KYLLFVFNFMVLAAGVILGVALMLRHPDQTSLLYLELGNKRPAPTFYVGIYILAVGA 68
QY 61 VMHVFVGFICGCAIQBSOCLGTFITFCVLIIIFACEVAAGIWGFVNKDQIAKDVQPYDQA 120
DB 69 LMMVLVGFICGCAIQBSOCLGTFITFCVLIIIFACEVAAGIWGFVNKDQIAKDVQPYDQA 128
QY 121 LQGVAMDDANNAKAVVKTFFHETLNCSSNALTTTLTLTLR--NSLCPSGNITLPLQQ 178
DB 129 YNKLKTRDPR--RETLLKAIHYALNCCG-----LAGVRFQSLSDICPK-KDVLFTFTVK 179
QY 179 DGHQKIDELFSGKLYLIGIAIVAVVIMIFEMILSMVLCGGIRNS 223

[illegible]

```
ch completed: December 9, 2003, 10:32:50
time : 38 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

on: December 9, 2003, 08:30:37 ; Search time 23 Seconds
(without alignments)

462.088 Million cell updates/sec

ect score: US-10-035-914-2

ence: 1 KYLLFVFNFWFLAGVILG.....IFEMILSWVLCGGIRNSSVY 226

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 127863 seqs, 47026705 residues

1 number of hits satisfying chosen parameters: 127863

mm DB seq length: 0

mm DB seq length: 200000000

-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

base: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

lc	Score	Query Match	Length	DB	ID	Description
1	1177	99.7	236	1	CD81_MOUSE	P35762 mus musculu
2	1120	94.9	236	1	CD81_RAT	O62745 rattus norv
3	1104	93.6	236	1	CD81_CERAE	O97703 cercoptihc
4	1103	93.5	236	1	CD81_HUMAN	P18582 homo sapien
5	1100	93.2	236	1	CD81_SAGOE	O90018 saguinu oe
6	528.5	44.8	225	1	CD9_BOVIN	P30933 bos taurus
7	525.5	44.5	225	1	CD9_FITG	O80643 sus scrofa
8	522.5	44.3	225	1	CD9_RAT	P40241 rattus norv
9	518.5	43.9	227	1	CD9_HUMAN	P41926 homo sapien
10	515.5	43.7	225	1	CD9_FELCA	P40239 felis silve
11	512.5	43.4	227	1	CD9_CERAE	P30409 cercoptihc
12	510.5	43.3	225	1	CD9_MOUSE	P40240 mus musculu
13	462.5	39.2	221	1	TSN2_RAT	O91741 rattus norv
14	458.5	38.9	221	1	TSN2_MOUSE	O92236 mus musculu
15	453.5	38.4	221	1	TSN2_HUMAN	O60636 homo sapien
16	298	25.3	237	1	T4S3_HUMAN	P19075 homo sapien
17	297.5	25.2	218	1	CD53_MOUSE	O61451 mus musculu
18	286.5	24.3	218	1	CD53_RAT	P24485 rattus norv
19	272.5	23.1	239	1	TNE5_HUMAN	O75954 homo sapien
20	272.5	23.1	267	1	CD82_HUMAN	P27701 homo sapien
21	265.5	22.5	219	1	CD83_HUMAN	P19397 homo sapien
22	262	22.2	266	1	CD82_MOUSE	P40237 mus musculu
23	254.5	21.6	238	1	T4S7_HUMAN	O14817 homo sapien
24	252.5	21.4	253	1	T4S7_MOUSE	O15566 mus musculu
25	250	21.2	266	1	CD81_MOUSE	O70352 rattus norv
26	248.5	21.1	266	1	CD81_RAT	O9q246 rattus norv
27	247.5	21.0	238	1	T4S7_MOUSE	O9dck3 mus musculu
28	244.5	20.7	253	1	C151_HUMAN	P48509 homo sapien
29	241.5	20.5	253	1	C151_CERAE	O9wmy2 cercoptihc
30	220.5	18.7	241	1	TSN1_HUMAN	O60635 homo sapien
31	209	17.7	249	1	T4S2_HUMAN	P41732 homo sapien
32	207	17.5	237	1	CD63_MOUSE	P41731 mus musculu
33	206	17.5	268	1	T4S9_HUMAN	O60628 homo sapien

ALIGNMENTS

RESULT 1	CD81_MOUSE	STANDARD;	PRT;	236 AA.
AC	P35762;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	CD81 antigen (26 kDa cell surface protein TAPA-1) (target of the antiproliferative antibody 1).			
GN	CD81 OR TAPAL.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91318144; PubMed=1650385;			
RA	Andria M.L., Hsieh C.L., Oren R., Francke U., Levy S.;			
RT	"Genomic organization and chromosomal localization of the TAPA-1 gene."			
RL	J. Immunol. 147:1030-1036(1991).			
RN	[2]			
RP	SEQUENCE OF 11-236 FROM N.A.			
RC	TISSUE=Heart;			
RA	Buff K., Parsons J.;			
RL	Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE REGULATION OF LYMPHOMA CELL GROWTH.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- PTM: NOT GLYCOSYLATED (PROBABLY).			
CC	-1- SIMILARITY: Belongs to the tetraspanin (TM4SF) family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; S45012; AAB19417.1; -			
DR	EMBL; S44957; AAB19417.1; JOINED.			
DR	EMBL; S44966; AAB19417.1; JOINED.			
DR	EMBL; S45001; AAB19417.1; JOINED.			
DR	EMBL; S45008; AAB19417.1; JOINED.			
DR	EMBL; S45010; AAB19417.1; JOINED.			
DR	EMBL; X59047; -; NOT_ANNOTATED_CDS.			
DR	PIR; A46472; A46472.			
DR	MCD; MGI:1096398; Cdbl.			
DR	InterPro: IPR000301; Transmem_4.			
DR	Pfam: PF00335; Transmembrane4; 1.			
DR	PRINTS; PR00259; TMPOUR.			
DR	PROSITE; PS00421; TM4_1; 1.			
KW	Transmembrane; Antigen.			
FT	DOMAIN 1 12			
FT	TRANSMEM 13 33			
				CYTOPLASMIC (POTENTIAL).
				POTENTIAL.

DOMAIN 34 57 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 58 78 POTENTIAL.
 DOMAIN 79 89 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 90 115 POTENTIAL.
 DOMAIN 116 201 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 202 226 POTENTIAL.
 DOMAIN 227 236 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 236 25828 MW; 7FCB386F79FA49 CRC64;
 SEQUENCE 236 AA; 25828 MW; 7FCB386F79FA49 CRC64;

CY Match 99.7%; Score 1177; DB 1; Length 236;
 Local Similarity 99.6%; Pred. No. 2.4e-93;
 Cons 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KYLLEFNFVFWLWAGVILGVALMLRHPDPTSLYLKGNKPAPNTFYVGIYLIANGA 60
 11 KYLLEFNFVFWLWAGVILGVALMLRHPDPTSLYLKGNKPAPNTFYVGIYLIANGA 70
 61 VMWFGVFCGCGAIQESQCLGTFPTCLVILPACVAVAGIWFVNDQIAKOVKQFYDOA 120
 71 VMWFGVFCGCGAIQESQCLGTFPTCLVILPACVAVAGIWFVNDQIAKOVKQFYDOA 130
 121 LQAAVMDDDANNAKAVVTFPHETLNCSSNALTTTLTILKNSLCPSSGNILTPLLQDDC 180
 131 LQAAVMDDDANNAKAVVTFPHETLNCSSNALTTTLTILKNSLCPSSGNILTPLLQDDC 190
 181 HOKIDELFSGKLYLIGIAIYVAVIMIFEMILSMWLCGGIRNSSYV 226
 191 HOKIDELFSGKLYLIGIAIYVAVIMIFEMILSMWLCGGIRNSSYV 236

1 2
 281 RAT STANDARD; PRT; 236 AA.
 362745;
 16-OCT-2001 (Rel. 40, Created)
 16-OCT-2001 (Rel. 40, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 D81 antigen (26 kDa cell surface protein TAPA-1) (Target of the
 antiproliferative antibody 1).
 D81 OR TAPA1.
 (actus norvegicus (Rat)).
 Burkayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lamellaria; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 [1] JCB1_TaxID=10116;

SEQUENCE FROM N.A.
 STRAIN=Sprague-Dawley; TISSUE=Brain cortex;
 MEDLINE=96346153; PubMed=8757260;
 JELSER B.B., Yang L., Irwin M.H.;
 Astrocyte growth, reactivity, and the target of the antiproliferative
 antibody, TAPA-1;
 Neurosci. 16:5478-5487 (1996).

1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE REGULATION OF LYMPHOMA
 CELL GROWTH.
 1- SUBCELLULAR LOCATION: Integral membrane protein.
 1- SIMILARITY: Belongs to the tetraspanin (TM4SF) family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

MBL; 019894; AAC3103.1;
 InterPro; IPR000301; Transmem 4.
 Pfam; PF00335; transmembrane4; 1.
 PRINTS; PR00259; TMPOUR.
 PROSITE; PS00421; TM4.1; 1.
 TRANSMEMbrane; Antigen.
 DOMAIN 1 12
 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 13 33 POTENTIAL.
 FT DOMAIN 34 57 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 58 78 POTENTIAL.
 FT DOMAIN 79 89 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 90 115 POTENTIAL.
 FT DOMAIN 116 201 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 202 226 POTENTIAL.
 FT DOMAIN 227 236 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 25888 MW; DCC48F38EB19BDP5 CRC64;
 SEQUENCE 236 AA; 25888 MW; DCC48F38EB19BDP5 CRC64;

Query Match 94.9%; Score 1120; DB 1; Length 236;
 Best Local Similarity 94.2%; Pred. No. 1.7e-88;
 Matches 213; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

1 KYLLEFNFVFWLWAGVILGVALMLRHPDPTSLYLKGNKPAPNTFYVGIYLIANGA 60
 11 KYLLEFNFVFWLWAGVILGVALMLRHPDPTSLYLKGNKPAPNTFYVGIYLIANGA 70
 61 VMWFGVFCGCGAIQESQCLGTFPTCLVILPACVAVAGIWFVNDQIAKOVKQFYDOA 120
 71 VMWFGVFCGCGAIQESQCLGTFPTCLVILPACVAVAGIWFVNDQIAKOVKQFYDOA 130
 121 LQAAVMDDDANNAKAVVTFPHETLNCSSNALTTTLTILKNSLCPSSGNILTPLLQDDC 180
 131 LQAAVMDDDANNAKAVVTFPHETLNCSSNALTTTLTILKNSLCPSSGNILTPLLQDDC 190
 181 HOKIDELFSGKLYLIGIAIYVAVIMIFEMILSMWLCGGIRNSSYV 226
 191 HOKIDELFSGKLYLIGIAIYVAVIMIFEMILSMWLCGGIRNSSYV 236

RESULT 3
 ID CD81_CERAB STANDARD; PRT; 236 AA.
 AC 097703;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE CD81 antigen.
 GN CD81.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecidae; Cercopithecus.
 OX NCBI_TaxID=9534;
 [1]

SEQUENCE FROM N.A.
 RA Levy S., Kuo C.C.;
 RT "African green monkey CD81 cDNA."
 RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
 CC 1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE REGULATION OF LYMPHOMA
 CELL GROWTH.
 CC 1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC 1- SIMILARITY: Belongs to the tetraspanin (TM4SF) family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

EMBL; AF116559; AAD11439.1;
 InterPro; IPR000301; Transmem 4.
 Pfam; PF00335; transmembrane4; 1.
 PRINTS; PR00259; TMPOUR.
 PROSITE; PS00421; TM4.1; 1.
 TRANSMEMbrane; Antigen.
 DOMAIN 1 12
 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 13 33 POTENTIAL.
 FT DOMAIN 34 57 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 58 78 POTENTIAL.

DOMAIN 79 89 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 90 115 POTENTIAL.
 DOMAIN 116 201 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 202 226 POTENTIAL.
 DOMAIN 227 236 CYTOPLASMIC (POTENTIAL).
 SEQUENCE 236 AA; 25758 MW; E07A8BC2C7637A4A CRC64;
 ery Match 93.6%; Score 1104; DB 1; Length 236;
 at Local Similarity 92.0%; Pred. No. 4e-87;
 tches 208; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

1 KTLFPNFPNFWLGGVITLGVALLMHRDQTTSLYLEGKRAEPTFYGIYLLAVGA 60
 |||||
 11 KTLFPNFPNFWLGGVITLGVALLMHRDQTTSLYLEGKRAEPTFYGIYLLAVGA 70
 |||||
 61 VMFVFGFLGCTGAIQGSQCLLGTFTCLVILFACVAAAGIAGVFNVDQIAKVKQFYDQA 120
 |||||
 71 VMFVFGFLGCTGAIQGSQCLLGTFTCLVILFACVAAAGIAGVFNVDQIAKVKQFYDQA 130
 |||||
 121 LQQAAMDDANNAAKAVKTFHEHTLNCSSNALTTTLTILRNSLCPGSGNITPLAQDC 180
 |||||
 131 LQQAAMDDANNAAKAVKTFHEHTLNCSSNALTTTLTILRNSLCPGSGNITPLAQDC 190
 |||||
 181 HOKIDELPSGKLYLIGIAIVAVVIMFEMILSMVLCGIRNSSYV 226
 |||||
 191 HOKIDELPSGKLYLIGIAIVAVVIMFEMILSMVLCGIRNSSYV 236
 |||||

LT 4
 HUMAN
 CD81 HUMAN STANDARD; PRT; 236 AA.
 P18582;
 01-NOV-1990 (Rel. 16, Created)
 01-NOV-1990 (Rel. 16, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 CD81 antigen (26 kDa cell surface protein TAP-1) (Target of the
 antiproliferative antibody 1).
 CD81 OR TAP-1.
 Homo sapiens (Human), and
 Pan troglodytes (Chimpanzee).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606, 9598;
 [1]
 SEQUENCE FROM N.A.
 SPECIES=Human;
 MEDLINE=90318365; PubMed=1695320;
 Oren R., Takahashi S., Doss C., Levy R., Levy S.,
 "TAP-1, the target of an antiproliferative antibody, defines a new
 family of transmembrane proteins.";
 Mol. Cell. Biol. 10:4007-4015(1990).
 [2]
 SEQUENCE FROM N.A.
 SPECIES=Human; TISSUE=Brain;
 MEDLINE=22388257; PubMed=12477932;
 Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Nak S.I., Wang J., Hsieh F.,
 Datchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
 Stepieton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Uedin T.B., Toshiyuki S., Carrinci P., Prange C.,
 Rana S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
 Boak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gnatrache P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Faye J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 Whitting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.,
 Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=P.troglodytes;
 RA Levy S., Kuo C.C.;
 RT "CD81 cDNA sequence derived from chimpanzee cells.";
 RN Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RP TOPOLOGY.
 RX MEDLINE=91317825; PubMed=1860863;
 RA Levy S., Nguyen V.O., Andria M.L., Takahashi S.;
 RT "Structure and membrane topology of TAP-1.";
 RN J. Biol. Chem. 266:14597-14602(1991).
 CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE REGULATION OF LYMPHOMA
 CELL GROWTH. INTERACTS WITH A 16-KDA LEU-13 PROTEIN TO FORM A
 COMPLEX POSSIBLY INVOLVED IN SIGNAL TRANSDUCTION. MAY ACTS A THE
 VIRAL RECEPTOR FOR HCV.
 CC -1- SUBUNIT: Interacts with Hepatitis C virus (HCV) glycoprotein E2.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: HEPATOLYMPHOID, NEUROECTODERMAL, & MESENCHYMAL
 TUMOR CELL LINES.
 CC -1- PTM: NOT GLYCOSYLATED (PROBABLE).
 CC -1- SIMILARITY: Belongs to the tetraspanin (TM4SF) family.
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD81 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd81.htm".

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 or send an email to license@ebi.ac.uk).

 DR EMBL; M33680; AAA36653.1; -;
 DR EMBL; BC002978; AAH02978.1; -;
 DR EMBL; AF116600; AA01440.1; -;
 DR PIR; A35649; A35649.
 DR PDB; 1G8Q; 21-FEB-01.
 DR PDB; 1IV5; 28-JAN-03.
 DR Genew; HGNC:1701; CD81.
 DR MIM; 186845; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR GO; GO:0006952; P:defense response; TAS.
 DR InterPro; IPR000301; Transmem 4.
 DR Pfam; PF00335; transmembrane4; 1.
 DR PRINTS; PR00259; TMFOUR.
 DR PROSITE; PS00421; TM4_1; 1.
 KW Transmembrane; Antigen; 3D-structure.
 FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 13 33 POTENTIAL.
 FT DOMAIN 34 63 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 64 84 POTENTIAL.
 FT DOMAIN 85 89 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 90 112 POTENTIAL.
 FT DOMAIN 113 201 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 202 224 POTENTIAL.
 FT DOMAIN 225 236 CYTOPLASMIC (POTENTIAL).
 SEQUENCE 236 AA; 25809 MW; EB9BD7671AC9184A CRC64;

Query Match 93.5%; Score 1103; DB 1; Length 236;
 Best Local Similarity 91.6%; Pred. No. 4.9e-87;
 Matches 207; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

OY 1 KTLFPNFPNFWLGGVITLGVALLMHRDQTTSLYLEGKRAEPTFYGIYLLAVGA 60
 |||||
 DB 11 KTLFPNFPNFWLGGVITLGVALLMHRDQTTSLYLEGKRAEPTFYGIYLLAVGA 70
 |||||
 OY 61 VMFVFGFLGCTGAIQGSQCLLGTFTCLVILFACVAAAGIAGVFNVDQIAKVKQFYDQA 120
 |||||
 DB 71 VMFVFGFLGCTGAIQGSQCLLGTFTCLVILFACVAAAGIAGVFNVDQIAKVKQFYDQA 130
 |||||

61 VMMVGFGLCCGAGVDSOCLLGTFTCLVILPACGVAAGIWFVNDQIAKDVQFYDQA 120
 66 LMMVGFGLCCGAGVDSOCLLGTFTCLVILPACGVAAGIWFVNDQIAKDVQFYDQA 125
 121 LQAAVDDDDANNAKAVVTFHETLNCCSGNALTTTLTTLNLSLCPSCGNILTPLLQDC 180
 126 YNKGKNDPQ--RRTLKAIHIALDCG--LTVGEQFLTYT-CPP-NLLIDSLKTRPC 178
 181 HOKIDELFSGKLYLIGIAIVAVIMIFEMILSMVLCGIRNS 223
 179 PEALKEVFNKHITIGAVGIGIAVIMIFEMILSMVLCGIRNS 221

LT 7
 CD9_PIG STANDARD; PRT; 225 AA.
 08MWQ3;
 28-FEB-2003 (Rel. 41, Created)
 28-FEB-2003 (Rel. 41, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 CD9 antigen.
 CD9.
 Sus scrofa (Pig).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9923;
 [1]
 SEQUENCE FROM N.A.
 Yubero N., Barbancho M.J., Llanes D., Garrido J.J.;
 "Molecular cloning of the pig homolog of tetraspanin CD9 antigen."
 Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 -1- FUNCTION: INVOLVED IN PLATELET ACTIVATION AND AGGREGATION (By
 similarity).
 -1- SUBCELLULAR LOCATION: Integral membrane protein.
 -1- SIMILARITY: Belongs to the tetraspanin (TM4SF) family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; AY072785; AAL68966.1; -
 InterPro: IPR000301; Transmem_4.
 Pfam: PF00335; Transmembrane4; 1.
 PRINTS; PR00259; TMFOUR.
 PROSITE; PS00421; TM4_1; 1.
 Glycoprotein; Antigen; Transmembrane; Lipoprotein.
 INIT MET 0 0
 DOMAIN 1 11
 TRANSSEM 12 32
 TRANSSEM 33 52
 TRANSSEM 53 73
 TRANSSEM 74 84
 TRANSSEM 85 108
 TRANSSEM 109 192
 TRANSSEM 193 218
 TRANSSEM 219 225
 CARBOHYD 49 49
 CARBOHYD 50 50
 SEQUENCE 225 AA; 25070 MW; FF280FB39BC1545 CRC64;

Query Match 44.5%; Score 525.5; DB 1; Length 225;
 Local Similarity 47.1%; Pred. No. 7.3e-38;
 Cons 105; Conservative 39; Mismatches 70; Indels 11; Gaps 5;
 1 KYLLFVNFVFMILAGVILGVALMLRHPDPTSLVLEIGKRPAPNFVYGIYLLAVGA 60
 10 KYLLFVNFVFMILAGVILGVALMLRHPDPTSLVLEIGKRPAPNFVYGIYLLAVGA 65
 61 VMMVGFGLCCGAGVDSOCLLGTFTCLVILPACGVAAGIWFVNDQIAKDVQFYDQA 120

Db
 66 LMMVGFGLCCGAGVDSOCLLGTFTCLVILPACGVAAGIWFVNDQIAKDVQFYDQA 125
 121 LQAAVDDDDANNAKAVVTFHETLNCCSGNALTTTLTTLNLSLCPSCGNILTPLLQDC 180
 126 YNKGKNDPQ--RRTLKAIHIALDCG--LTVGEQFLTYT-CPP-NLLIDSLKTRPC 178
 181 HOKIDELFSGKLYLIGIAIVAVIMIFEMILSMVLCGIRNS 223
 179 PEALKEVFNKHITIGAVGIGIAVIMIFEMILSMVLCGIRNS 221

RESULT 8
 CD9_RAT STANDARD; PRT; 225 AA.
 ID CD9_RAT
 AC P40241;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE CD9 antigen.
 CN
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Sciatic nerve;
 RX MEDLINE=95123481; PubMed=7823164;
 RA Kaprielian Z., Cho K.O., Hadjilargyrou M., Patterson P.H.;
 RT "CD9, a major platelet cell surface glycoprotein, is a ROCA antigen
 and is expressed in the nervous system."
 RL J. Neurosci. 15:562-573 (1995).
 CC -1- FUNCTION: INVOLVED IN PLATELET ACTIVATION AND AGGREGATION
 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE PERIPHERAL NERVOUS SYSTEM.
 CC -1- SIMILARITY: Belongs to the tetraspanin (TM4SF) family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; X76489; CA54027.1; -
 DR PIR: I56562; S39262.
 DR InterPro: IPR000301; Transmem_4.
 DR Pfam: PF00335; Transmembrane4; 1.
 DR PRINTS; PR00259; TMFOUR.
 DR PROSITE; PS00421; TM4_1; 1.
 KM Glycoprotein; Antigen; Transmembrane.
 FT INIT MET 0 0
 FT DOMAIN 1 11
 FT TRANSSEM 12 32
 FT TRANSSEM 33 52
 FT TRANSSEM 53 73
 FT TRANSSEM 74 84
 FT TRANSSEM 85 108
 FT TRANSSEM 109 192
 FT TRANSSEM 193 218
 FT TRANSSEM 219 225
 FT CARBOHYD 49 49
 FT CARBOHYD 50 50
 SEQUENCE 225 AA; 25084 MW; 7889619P99A59C9D CRC64;

Query Match 44.3%; Score 522.5; DB 1; Length 225;
 Best Local Similarity 45.1%; Pred. No. 1.3e-37;
 Matches 101; Conservative 39; Mismatches 71; Indels 13; Gaps 5;
 1 KYLLFVNFVFMILAGVILGVALMLRHPDPTSLVLEIGKRPAPNFVYGIYLLAVGA 60
 61 VMMVGFGLCCGAGVDSOCLLGTFTCLVILPACGVAAGIWFVNDQIAKDVQFYDQA 120

10 KYLLFGNFIFLMIAGIIVLIGLWLRDPDSOTKS1PEOBTHN-----SFFYGVILLIGAGA 65
 11 VMFVFGYGCYGAIOESOCILGFEPTGVLIFACEVAGIWMGFVNKOIKNDVQGFYDOA 120
 12 LMLVGLTGGCGAVQESQCHLGFELVIFALFIAAAWGYTHKDEVIKEIQEFTYKT 125
 121 LQQAAMDANANAAVAVTPEHTLNCCG-SNALTTTLTTLTILNLSLCPSGGNILPLPLQD 179
 126 YOKLRNDEPQ-RETLKALHMLNCCGINGVGFISDI-----CK-KQVLESPOVKS 177
 180 CHQKIDELFSGKYLITGIAIIVAVNIFEMILSMVLCCGIRNS 223
 178 CPDAIDEVFHSKFIHIGAVGIGIIVMIFGMIFSMILCAIRRS 221

T 9
 UMAN
 CD9 HUMAN STANDARD; PRT; 227 AA.
 P21926; Q96RS4;
 01-MAY-1991 (Rel. 18, Created)
 01-APR-1993 (Rel. 25, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 CD9 antigen (P24) (Leukocyte antigen M1C3) (Motility-related protein) (MRP-1).
 CD9 OR M1C3.
 Homo sapiens (Human).
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_TaxID=9606;
 (1)
 SEQUENCE FROM N.A., AND SEQUENCE OF 1-4.
 MEDLINE=91093112; PubMed=1840589;
 Boucheix C., Benoit P., Prachet P., Billard M., Worthington R.B., Zagon J., Uzan G.;
 "Molecular cloning of the CD9 antigen. A new family of cell surface proteins.";
 J. Biol. Chem. 266:117-122(1991).
 (2)
 SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 MEDLINE=9124846; PubMed=2037607;
 Janza F., Wolf D., Fox C.F., Kieffer N., Seyer J.M., Fried V.A., Joughlin S.R., Phillips D.R., Jennings L.K.;
 "cDNA cloning and expression of platelet P24/CD9. Evidence for a new family of multiple membrane-spanning proteins.";
 J. Biol. Chem. 266:10638-10645(1991).
 (3)
 SEQUENCE FROM N.A.
 MEDLINE=92078843; PubMed=1720807;
 Miyake M., Koyama M., Seno M., Ikegama S.;
 "Identification of the motility-related protein (MRP-1), recognized by monoclonal antibody M31-15, which inhibits cell motility.";
 J. Exp. Med. 174:1347-1354(1991).
 (4)
 SEQUENCE FROM N.A.
 TISSUE=Leukocyte;
 MEDLINE=93252369; PubMed=8466348;
 Rubinstejn B., Benoit P., Billard M., Plaisance S., Prenant M., Uzan G., Boucheix C.;
 "Organization of the human CD9 gene.";
 Genomics 16:132-138(1993).
 (5)
 SEQUENCE FROM N.A.
 TISSUE=Ovary;
 MEDLINE=22388257; PubMed=12477932;
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., White S.P., Zeeberg B., Bueltow K.H., Scheafer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Jatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stajich M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.R., Jahn S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullen S.J., Josak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Rahay J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schenker A., Schein J.E., Jones S.J.M., Maita M.A.,
 RA "regeneration and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP SEQUENCE OF 1-20.
 RC TISSUE=Platelet;
 RA MEDLINE=90292223; PubMed=2358073;
 RA Higashihara M., Takahata K., Yatom Y., Nakahara K., Kurokawa K.;
 RT "Purification and partial characterization of CD9 antigen of human
 RT platelets.";
 RL FEBS Lett. 264:270-274(1990).
 RN [7]
 RP PHOSPHORYLATION.
 RA MEDLINE=93327758; PubMed=7687539;
 RA Chalupny N.J., Kanner S.B., Schieven G.L., Wee S., Gilliland L.K.,
 RA Arafio A., Ledbetter J.A.;
 RT "Tyrosine phosphorylation of CD19 in pre-B and mature B cells.";
 RL EMBO J. 12:2691-2696(1993).
 RN [8]
 RP INTERACTION WITH PTGFRN.
 RA MEDLINE=21216740; PubMed=11278880;
 RA Charrin S., Le Naour F., Oualid M., Billard M., Faure G., Hanash S.M.,
 RA Boucheix C., Rubinstein E.;
 RT "The major CD9 and CD81 molecular partner. Identification and
 RT characterization of the complexes.";
 RL J. Biol. Chem. 276:14329-14337(2001).
 CC -1- FUNCTION: INVOLVED IN PLATELET ACTIVATION AND AGGREGATION.
 CC -1- SUBUNIT: Associates with CR2/CD21 and with PTGFRN/CD91.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED BY A VARIETY OF HEMATOPOIETIC
 CC AND EPITHELIAL CELLS.
 CC -1- PTM: PROTEIN EXISTS IN THREE FORMS WITH MOLECULAR MASSES BETWEEN
 CC 22 AND 27 kDa, AND IS KNOWN TO CARRY COVALENTLY LINKED FATTY
 CC ACIDS.
 CC -1- PTM: PHOSPHORYLATED ON TYROSINE FOLLOWING B CELL ACTIVATION.
 CC -1- SIMILARITY: Belongs to the tetraspanin (TM4SF) family.
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD9 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd9.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL, S60489; AAC60586.1; JOINED.
 DR EMBL, S60462; AAC60586.1; JOINED.
 DR EMBL, S60463; AAC60586.1; JOINED.
 DR EMBL, S60464; AAC60586.1; JOINED.
 DR EMBL, S60700; AAC60586.1; JOINED.
 DR EMBL, S60699; AAC60586.1; JOINED.
 DR EMBL, S60465; AAC60586.1; JOINED.
 DR EMBL, S60472; AAC60586.1; JOINED.
 DR EMBL, M38690; AAA59382.1; -.
 DR EMBL, L34068; AAA59382.1; -.
 DR EMBL, X60111; CAA42708.1; -.
 DR EMBL, L08118; -; NOT ANNOTATED CDS.
 DR EMBL, L08119; AAA51954.1; ALT_SEQ.
 DR EMBL, L08120; AAA51955.1; ALT_SEQ.
 DR EMBL, L08121; AAA51956.1; -.
 DR EMBL, L08122; AAA51957.1; -.
 DR EMBL, L08123; AAA51958.1; -.
 DR EMBL, L08124; AAA51959.1; -.

[illegible]

vero cells is the monkey homologue of human CD9 antigen: expression of DRAP27 elevates the number of diphtheria toxin receptors on toxin-sensitive cells."

J. Cell Biol. 118:1389-1399(1992).

-1- FUNCTION: EXPRESSION OF DRAP27 ELEVATES THE NUMBER OF DIPHTHERIA TOXIN RECEPTORS ON TOXIN-SENSITIVE CELLS.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- SIMILARITY: Belongs to the tetraspanin (TM4SF) family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D10726; BA01569.1; --

PIR; A42929; A42929.

InterPro; IPR000301; Transmem 4.

Pfam; PF00335; transmembrane4; 1.

PRINTS; PR00259; TMFOUR.

PROSITE; PS00421; TM4_1; 1.

INIT MET 0 BY SIMILARITY.

DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).

TRANSSEM 12 32 POTENTIAL.

DOMAIN 33 54 EXTRACELLULAR (POTENTIAL).

TRANSSEM 55 75 POTENTIAL.

DOMAIN 76 86 CYTOPLASMIC (POTENTIAL).

TRANSSEM 87 110 POTENTIAL.

DOMAIN 111 194 EXTRACELLULAR (POTENTIAL).

TRANSSEM 195 220 POTENTIAL.

DOMAIN 221 227 CYTOPLASMIC (POTENTIAL).

CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).

227 AA; 25300 MW; 8834748759374A CRC64;

Query Match 43.4%; Score 512.5; DB 1; Length 227;

Local Similarity 44.4%; Pred. No. 9.4e-37;

Ches 100; Conservative 40; Mismatches 72; Indels 13; Gaps 5;

1 KYLLFVFNFMVLGAVILGVALMLRHPDPTSLYLKGNKPAFTFYGIYLLAVGA 60

10 KYLLFVFNFMVLGAVILGVALMLRHPDPTSLYLKGNKPAFTFYGIYLLAVGA 67

61 VMVVFGLGCTGAIQESQCLGTFPTCLVILPACVAVAGIWMGVNKKQIAKDYKQFYDQ 120

68 LMMLVGFLGCGAVQESQCLGTFPTCLVILPACVAVAGIWMGVNKKQIAKDYKQFYDQ 127

121 LQAVVDDANNAKAVVKTFFHTLNCGGSNALTTTLTIIR--NSLCPSGNTLTPLLQD 178

128 YNKAKTKDQBPQ--RETLKALHVALDCCG-----LACGVQFISDICKP-KDVALETTIK 178

179 DCHOKIDELFSKGLYLIGIAIYVAVIMIFEMILSMVLCGGIRNS 223

179 SCDDAIKEVDFDNKFIIGAVGIGAVMIFGMIFSMILCCAIRN 223

1 12

CD9 MOUSR STANDARD; PRT; 225 AA.

PA0240;

01-FEB-1995 (Rel. 31, Created)

01-FEB-1995 (Rel. 31, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)

CD9 antigen.

CD9.

Mus musculus (Mouse).

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxId=10090;

[1]

SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE=Kidney;

RX MEDLINE-94054345; PubMed=8236164;

RA Rudinstein E., Billard M., Plaisance S., Prenant M., Boucheix C.;

RT Molecular cloning of the mouse equivalent of CD9 antigen."

Thromb. Res. 71:377-383(1993).

-1- FUNCTION: INVOLVED IN PLATELET ACTIVATION AND AGGREGATION

(BY SIMILARITY).

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- SIMILARITY: Belongs to the tetraspanin (TM4SF) family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; I08115; AAA37405.1; --

PIR; I49589; I49589.

MGI; MGI:88348; CD9.

InterPro; IPR000301; Transmem 4.

Pfam; PF00335; transmembrane4; 1.

PRINTS; PR00259; TMFOUR.

PROSITE; PS00421; TM4_1; 1.

INIT MET 0 BY SIMILARITY.

DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).

TRANSSEM 12 32 POTENTIAL.

DOMAIN 33 52 EXTRACELLULAR (POTENTIAL).

TRANSSEM 53 73 POTENTIAL.

DOMAIN 74 84 CYTOPLASMIC (POTENTIAL).

TRANSSEM 85 108 POTENTIAL.

DOMAIN 109 192 EXTRACELLULAR (POTENTIAL).

TRANSSEM 193 218 POTENTIAL.

DOMAIN 219 225 CYTOPLASMIC (POTENTIAL).

CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 225 AA; 25127 MW; 2BB840B8D7C31BC0 CRC64;

Query Match 43.3%; Score 510.5; DB 1; Length 225;

Best Local Similarity 44.2%; Pred. No. 1.4e-36;

Matches 99; Conservative 41; Mismatches 71; Indels 13; Gaps 5;

1 KYLLFVFNFMVLGAVILGVALMLRHPDPTSLYLKGNKPAFTFYGIYLLAVGA 60

10 KYLLFVFNFMVLGAVILGVALMLRHPDPTSLYLKGNKPAFTFYGIYLLAVGA 65

61 VMVVFGLGCTGAIQESQCLGTFPTCLVILPACVAVAGIWMGVNKKQIAKDYKQFYDQ 120

66 LMMLVGFLGCGAVQESQCLGTFPTCLVILPACVAVAGIWMGVNKKQIAKDYKQFYDQ 125

121 LQAVVDDANNAKAVVKTFFHTLNCGG--SNALTTTLTIIRNSLCPSGNTLTPLLQD 179

126 YQKRSKQBPQ--RETLKALHVALDCCG--GAGLEGP-----SDTCPK-KQLLESQVNP 177

180 DCHOKIDELFSKGLYLIGIAIYVAVIMIFEMILSMVLCGGIRNS 223

178 CPDAISEVENKFIIGAVGIGAVMIFGMIFSMILCCAIRS 221

RESULT 13

TSN2_RAT

ID TSN2_RAT STANDARD; PRT; 221 AA.

AC 09JUM1;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

GN Tetraspanin 2 (Tspan-2).

OS TSPAN2.

OS Rattus norvegicus (Rat).

OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

Protein - protein search, using sw model -

on: December 9, 2003, 09:35:04 / Search time 85 Seconds

(Without alignments)
686.116 Million cell updates/sec

ect score: 1180

ence: 1 KTLFVFNFWLACGVILG.....IFEMILSMVLCCGIRNSVY 226

ing table: BLOSUM62

ched: 830525 seqs, 258052604 residues

1 number of hits satisfying chosen parameters: 830525

num DB seq length: 0

num DB seq length: 2000000000

-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_prodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_ryins:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query Match Length DB ID Description

Query	Match	Length	DB	ID	Description
1	1180	100.0	236	11	Q91V78 mus musculus
2	1018	86.3	211	16	Q9GLG9
3	738	62.5	236	13	Q9DD57
4	616.5	52.2	228	13	Q8AV92
5	523.5	44.4	224	13	Q9IBG9
6	519.5	44.0	226	6	Q8M748
7	486	41.2	222	13	Q8AV91
8	422.5	35.8	199	11	Q8COC9
9	411.5	34.9	217	11	Q9D1X8
10	300	25.4	235	11	Q8R3G9
11	297	25.2	235	11	Q8R3G9
12	290.5	24.6	268	5	Q9V3X2
13	288	24.4	248	4	Q96S78
14	283.5	24.0	247	13	Q9PTB0
15	274.5	23.3	239	11	Q8BTU2
16	272	23.1	236	4	Q8NB19

Query	Match	Length	DB	ID	Description
17	269.5	22.8	230	13	Q8AVJ1
18	267.5	22.7	267	5	Q9V3E5
19	260.5	22.1	253	11	Q9D1D1
20	259	21.9	267	4	Q8WUV1
21	257	21.8	240	11	Q9V3J9
22	255.5	21.7	282	5	Q19983
23	251.5	21.3	253	11	Q921J7
24	242	20.5	283	11	Q8R3S2
25	239.5	20.3	230	13	Q8AV00
26	234	19.8	270	4	Q9H0U1
27	231	19.6	239	5	Q9NB10
28	231	19.6	270	4	Q8WV11
29	230	19.5	270	11	Q8QZV6
30	227.5	19.3	233	5	Q18177
31	220	18.6	270	11	Q91V16
32	220	18.6	270	11	Q9D7M4
33	219.5	18.6	236	5	Q9VVM5
34	218.5	18.5	223	5	Q9U2K6
35	211.5	17.9	227	5	Q8T9S4
36	211	17.9	232	5	Q9NB07
37	210.5	17.8	232	5	Q9V174
38	207	17.5	222	5	Q917D5
39	207	17.5	249	11	Q8BT06
40	206	17.5	266	11	Q9CTY0
41	204	17.3	239	5	Q8MKY8
42	202	17.1	263	4	Q96PV3
43	201.5	17.1	276	5	Q17526
44	199.5	16.9	244	5	Q46101
45	199	16.9	291	5	Q9VGV3

ALIGNMENTS

RESULT 1

Q91V78 PRELIMINARY; PRT; 236 AA.

AC Q91V78 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Similar to CD81 antigen (Target of antiproliferative antibody 1)

DE (Tapa-1 protein)

GN TREM5 OR CD81 OR TAPA-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10990;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Breast tumor;

RA Strauberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RM [2]

RP SEQUENCE FROM N.A.

RC STRAIN=129 SV;

RK MEDLINE=20377495; PubMed=10915772;

RA Paulsen M., El-Maari O., Engemann S., Stroedicke M., Franck O.,

RA Davies K., Reinhardt R., Reik W., Walter J.;

RT "Sequence conservation and variability of imprinting in the Beckwith-

RT Wiedemann syndrome gene cluster in human and mouse.";

RL Hum. Mol. Genet. 9:1829-1841(2000).

DR EMBL; BC011433; AAH11433.1; -.

DR EMBL; AJ251835; CAB94774.1; -.

DR MGD; MGI:1096398; Cd81.

DR MGD; MGI:1861718; Trpm5.

DR InterPro; IPR000301; Transmem_4.

DR PRINTS; PR00259; TMFOUR.

DR PROSITE; PS00421; TM4_1; 1.

SO SEQUENCE 236 AA; 25814 MW; 7BDCB792E79FAAE CRC64;

Query Match 100.0%; Score 1180; DB 11; Length 236;

Best Local Similarity 100.0%; Pred. No. 5.2e-105;

ches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KYLLFVNFVFWLACGVIIIGVALMLRHDPTTSLYLKGNKPAKPTFYGYIILAVGA 60
 11 KYLLFVNFVFWLACGVIIIGVALMLRHDPTTSLYLKGNKPAKPTFYGYIILAVGA 70
 61 VMVFGVFCYGAIOESQCLGTFPTCLVILFACVAAIGWPNKQIAKDKOPYDQA 120
 71 VMVFGVFCYGAIOESQCLGTFPTCLVILFACVAAIGWPNKQIAKDKOPYDQA 130
 121 LQAAVDDDDANNAKAVKTFHETLNCSSNALTTTLTILRNSLCPSGNLTPLLOQDC 180
 131 LQAAVDDDDANNAKAVKTFHETLNCSSNALTTTLTILRNSLCPSGNLTPLLOQDC 190
 181 HKKIDELFSKGLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 226
 191 HKKIDELFSKGLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 236

T 2
 9
 09GLG9 PRELIMINARY; PRT; 211 AA.
 01-MAR-2001 (T-EMBLrel. 16, Created)
 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
 CD81 (Fragment).
 Saguinus oedipus (Cotton-top tamarin).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Placental; Platyrrhini; Callitrichidae; Saguinus.
 NCBI_TaxID=9490;

SEQUENCE FROM N.A.
 MEDLINE=20534982; PubMed=11080483;
 Allander T., Fornis X., Emerson S.U., Purcell R.H., Bukh J.;
 "Hepatitis C virus envelope protein E2 binds to CD81 of tamarins.";
 Virology 277:358-367(2000)
 EMBL; AF274885; AAC31990.1; -
 InterPro; IPR000301; Transmem_4.
 Pfam; PF00335; transmembrane4; 1.
 PRINTS; PR00259; TMFOUR.
 PROSITE; PS00421; TM_1; 1.
 NON_TER 1 211
 SEQUENCE 211 AA; 23091 MW; CACB34E89DAD3BD CRC64;

CY Match 86.3%; Score 1018; DB 6; Length 211;
 Local Similarity 91.0%; Pred. No. 1.5e-89;
 ches 192; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

4 LRVFNVFWLACGVIIIGVALMLRHDPTTSLYLKGNKPAKPTFYGYIILAVGA 63
 1 LRVFNVFWLACGVIIIGVALMLRHDPTTSLYLKGNKPAKPTFYGYIILAVGA 60
 64 PVGFGVFCYGAIOESQCLGTFPTCLVILFACVAAIGWPNKQIAKDKOPYDQA 123
 61 PVGFGVFCYGAIOESQCLGTFPTCLVILFACVAAIGWPNKQIAKDKOPYDQA 120
 124 AVMDDDANNAKAVKTFHETLNCSSNALTTTLTILRNSLCPSGNLTPLLOQDC 183
 121 AVMDDDANNAKAVKTFHETLNCSSNALTTTLTILRNSLCPSGNLTPLLOQDC 180
 184 IDELFSKGLYLIGIAIVAVVIMIFEMILSM 214
 181 IDELFSKGLYLIGIAIVAVVIMIFEMILSM 211

F 3
 7
 39DSD7 PRELIMINARY; PRT; 236 AA.
 39DSD7; (T-EMBLrel. 16, Created)
 31-MAR-2001 (T-EMBLrel. 16, Last sequence update)

DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE CD81.
 GN CD81.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21100417; PubMed=11167010;
 RA Yoder J.A., Litman G.W.;
 RT "The zebrafish fish, slc3a2, men1, pc, fgf3 and cycd1 genes define two
 RT regions of conserved synteny between linkage group 7 and human
 RT chromosome 11q13.";
 RL Gene 261:235-242(2000).
 DR EMBL; AF295377; AAC37840.1; -
 DR ZFIN; ZDB-GENE-000831-5; cd81.
 DR InterPro; IPR001092; HLH_Basic.
 DR Pfam; PF00335; transmembrane4; 1.
 DR PRINTS; PR00259; TMFOUR.
 DR PROSITE; PS00038; HLH_1; 1.
 DR PROSITE; PS00421; TM_1; 1.
 SQ SEQUENCE 236 AA; 26168 MW; 3C9964B8FD953428 CRC64;

Query Match 62.5%; Score 738; DB 13; Length 236;
 Best Local Similarity 64.0%; Pred. No. 1.1e-62;
 Matches 146; Conservative 24; Mismatches 52; Indels 6; Gaps 3;

QY 1 KYLLFVNFVFWLACGVIIIGVALMLRHDPTTSLYLKGNKPAKPTFYGYIILAVGA 60
 DB 13 KYLLFVNFVFWLACGVIIIGVALMLRHDPTTSLYLKGNKPAKPTFYGYIILAVGA 72
 QY 61 VMVFGVFCYGAIOESQCLGTFPTCLVILFACVAAIGWPNKQIAKDKOPYDQA 120
 DB 73 VMVFGVFCYGAIOESQCLGTFPTCLVILFACVAAIGWPNKQIAKDKOPYDQA 132
 QY 121 LQAAVDDDDANNAKAVKTFHETLNCSSNALTTTLTILRNSLCPSGNLTPLLOQDC 178
 DB 133 YKGAATYVNDNPNPATVILKVFHETLNCSSNALTTTLTILRNSLCPSGNLTPLLOQDC 188
 QY 179 DCHOKIDELFSKGLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 226
 DB 189 DCHOKIDELFSKGLYLIGIAIVAVVIMIFEMILSMVLCGGIRNSSVY 236

RESULT 4
 Q8AV92 PRELIMINARY; PRT; 228 AA.
 AC Q8AV92;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE CD9-like protein.
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 NCBI_TaxID=7557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymusole;
 RX MEDLINE=22295056; PubMed=12391333;
 RA Ulnak-Ool T., Mayer W.R., Sato A., Dongak R., Cooper M.D., Klein J.;
 RT "Lamprey lymphocyte-like cells express homologs of genes involved in
 RT immunologically relevant activities of mammalian lymphocytes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14356-14361(2002).
 DR EMBL; AY152677; AAN64299.1; -
 SQ SEQUENCE 228 AA; 25331 MW; DDE0EA31B988935 CRC64;

Query Match 52.2%; Score 616.5; DB 13; Length 228;
 Best Local Similarity 51.3%; Pred. No. 4.7e-51;
 Matches 116; Conservative 45; Mismatches 56; Indels 9; Gaps 5;

1 KTLFVNFVFWLGGVILGVALMLRHDPTSLYLTLKPKAPNTFYGIYILAVGA 60
 11 KTLFVNFVFWLGGVILGVALMLRHDPTSLYLTLKPKAPNTFYGIYILAVGA 66
 61 VMMFVGLGCGAIOESQCLGTFCTCLVILFACVAAAGIWFVFNKQIADVKQFYDQA 120
 67 LMMVGLGCGAIOESQCLGTFCTCLVILFACVAAAGIWFVFNKQIADVKQFYDQA 125
 121 LQOAVMDDDANNAKAVKTFHETLNCSSNALTTTLTLNLSLCPSSGNILTPLODC 180
 126 LVHEHKKTDSSKALITFHNTLSCCGMAYTNLSDSPKO-ICPT-OKITDPTV-DC 181
 181 HOKIDELFSGKLYIGIAIIVAVIMIFEMILSMVLCGIRNSVY 226
 182 IROIDEIFTSKYIACIGIGIVIMIFGMIFSMALCCAVRNRRDY 227

LT 5

091BC9 PRELIMINARY; PRT; 224 AA.

091BC9; 01-OCT-2000 (Tremblrel. 15, Created)
 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 CD9 antigen.

CD9.
 Gallus gallus (Chicken).

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 Gallus.

NCBI_TaxID=9031;

[1]
 SEQUENCE FROM N.A.

TISSUE=Spinal cord;
 Kobayashi T., Mekada E.;
 "Molecular cloning and characterization of Chick CD9,"
 Submitted (SBP-1999) to the EMBL/Genbank/DBJ databases.
 EMBL; AB032767; BAA3011.1; -
 InterPro: IPR000301; Trasnemem_4.
 Pfam: PF00335; transmembrane4; 1.
 PRINTS; PR00259; TMFOUR.
 PROSITE; PS00421; TM4_1; 1.

SEQUENCE 224 AA; 24969 MW; AB64BDF849516151 CRC64;

Query Match 44.4%; Score 523.5; DB 13; Length 224;
 Best Local Similarity 47.3%; Pred. No. 3.7e-42;
 Matches 107; Conservative 31; Mismatches 69; Indels 19; Gaps 5;

1 KTLFVNFVFWLGGVILGVALMLRHDPTSLYLTLKPKAPNTFYGIYILAVGA 60
 11 KTLFVNFVFWLGGVILGVALMLRHDPTSLYLTLKPKAPNTFYGIYILAVGA 66
 61 VMMFVGLGCGAIOESQCLGTFCTCLVILFACVAAAGIWFVFNKQIADVKQFYDQA 120
 67 LMMVGLGCGAIOESQCLGTFCTCLVILFACVAAAGIWFVFNKQIADVKQFYDQA 126
 121 LQOAVMDDDANNAKAVKTFHETLNCSSNALTTTLTLNLSLCPSSGNILTPLODC 177
 127 YEKK---SOPAAEILKAFHFAINCCG-----ITGGLEOQIMESCPK-KTLPSFTT 174
 178 QDCHQKIDELFSGKLYIGIAIIVAVIMIFEMILSMVLCGIRNS 223
 175 MPEPKAIDVFNSTLVNIGAVGIAIVIMIFGMIFSMALCCAVRN 220

88MJ48 PRELIMINARY; PRT; 226 AA.
 88MJ48; 01-OCT-2002 (Tremblrel. 22, Created)
 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Transmembrane protein CD9.
 CN CD9.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Heinz M., Huang C.A., Emery D.W., Giovino M.A., LeGuern A.,
 RA Kurilla-Mahon B., Theodore P., Arn J.S., Sykes M., Mulligan R.C.,
 RA Down J.D., Sachs D.H., Godeill M.A.;
 RT "Use of CD9 Expression to Enrich for Porcine Hematopoietic
 RT Progenitors."
 RT Exp Hematol. 0:0-0(2002).
 DR EMBL; AF525029; AAB81376.1; -
 DR InterPro; IPR000301; Trasnemem_4.
 DR PRINTS; PR00259; TMFOUR.
 DR PROSITE; PS00421; TM4_1; 1.
 SQ SEQUENCE 226 AA; 25127 MW; 885C244477F95C23 CRC64;

Query Match 44.0%; Score 519.5; DB 6; Length 226;
 Best Local Similarity 46.6%; Pred. No. 9.1e-42;
 Matches 104; Conservative 37; Mismatches 71; Indels 11; Gaps 5;

QY 1 KTLFVNFVFWLGGVILGVALMLRHDPTSLYLTLKPKAPNTFYGIYILAVGA 60
 DB 11 KTLFVNFVFWLGGVILGVALMLRHDPTSLYLTLKPKAPNTFYGIYILAVGA 66
 QY 61 VMMFVGLGCGAIOESQCLGTFCTCLVILFACVAAAGIWFVFNKQIADVKQFYDQA 120
 DB 67 LMMVGLGCGAIOESQCLGTFCTCLVILFACVAAAGIWFVFNKQIADVKQFYDQA 126
 QY 121 LQOAVMDDDANNAKAVKTFHETLNCSSNALTTTLTLNLSLCPSSGNILTPLODC 180
 DB 127 YNKKKGDDPQ--RETUKAHVYALDCCG---LMGVSQQLAD-ICPO-KDVLSSLPKPC 179
 QY 181 HOKIDELFSGKLYIGIAIIVAVIMIFEMILSMVLCGIRNS 223
 DB 180 PEALKEVFNKFNHIGIIVGIAIVIMIFGMIFSMALCCAVRN 222

RESULT 7

QY 08AV91 PRELIMINARY; PRT; 222 AA.

AC 08AV91; 01-MAR-2003 (Tremblrel. 23, Created).
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DB CD9-like protein.
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=lythiosole;
 RX MEDLINE=2295056; Pubmed=12391333;
 RA Ulunk-Ool T., Mayer W.E., Sato A., Dongak R., Cooper M.D., Klein J.;
 RT "Lamprey lymphocyte-like cells express homologs of genes involved in
 RT immunologically relevant activities of mammalian lymphocytes."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14356-14361(2002).
 DR EMBL; AY152678; AAB64300.1; -
 SQ SEQUENCE 222 AA; 24637 MW; CAC0A561301EFB54 CRC64;

Query Match 41.2%; Score 486; DB 13; Length 222;
 Best Local Similarity 43.5%; Pred. No. 1.4e-38;
 Matches 100; Conservative 40; Mismatches 68; Indels 22; Gaps 6;

QY 1 KTLFVNFVFWLGGVILGVALMLRHDPTSLYLTLKPKAPNTFYGIYILAVGA 60
 DB 11 KTLFVNFVFWLGGVILGVALMLRHDPTSLYLTLKPKAPNTFYGIYILAVGA 66
 QY 61 VMMFVGLGCGAIOESQCLGTFCTCLVILFACVAAAGIWFVFNKQIADVKQFYDQA 120

```

67 1MMVGGGLGCGCAIQBSQCMGTFETLIVILFSEIDRAVAVAFHDEVEVETQMLEYNN--124
121 LQAVVDDDDNNK-AVVT---FHETLNCSSALTTTLTILNSLCCSGGNILTPLL 176
125 ---GLINDNSNSTRSASKTSLIHSITLDCGCPNML--YTSILSTYSCP-----PEY 172
177 QODCHOKIDELFGSKYLYLIGIAIVAVIMIFEMILSMVLCGGIRNSSY 226
173 TKKCVDEIRSFGLTKYIYVAGLAVAVIMIFGMIFMAVCAIRNDSPY 222

7 8
9 PRELIMINARY; PRT; 199 AA.
28C0C9;
01-MAR-2003 (TREMBLrel. 23, Created)
01-MAR-2003 (TREMBLrel. 23, Last sequence update)
01-MAR-2003 (TREMBLrel. 23, Last annotation update)
TSPAN-2 protein homolog (Fragment).
Mus musculus (Mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
3RAIN=C57BL/6J; TISSUE=Testis;
MEDLINE=22354683; PubMed=12466851;
The FANTOM Consortium,
the RIKEN Genome Exploration Research Group Phase I & II Team;
"Analysis of the mouse transcriptome based on functional annotation of
50,770 full-length cDNAs.";
Nature 420:563-573(2002).
EMBL; AK031683; BAC27511.1; -.
NON TER 1
3SEQUENCE 199 AA; 21611 MW; 5DD1929442680978 CRC64;

CY Match 35.8%; Score 422.5; DB 11; Length 199;
Local Similarity 40.0%; Pred. No. 1.6e-32;
Indels 23; Gaps 5;
Conservative 31; Mismatches 75;

12 WTAGVILGVALMLRHPDPTSLIYELGKRPAPNTFYGYIYLIVAGVMMVFGGCG 71
1 WTAGSAVIAFGMLFRGGTMKDL-----SSEDKSPRYFYGLVGVAGALMMVFGGCG 56
72 GAIOBQCLLGTFTCLVILFACVAAAGIAGVFNKQDIADVQFTDQALQAVDDDDAN 131
57 GMRBESQCVLGSFFTCILVIFAAVTVGVFAFGKQVAIRHVSMEAVSDYLKDRARG 116
132 NAAVAVKTHETLNCSSNALTTTLTILNSLCPG---GNNLTPLLODCHOKIDELF 188
117 NGTLI--THSAFQCCGKSSSQVOPT-----CPKELPG-----HKNCIDKLETVA 160

189 SGKLYLIGIAIVAVIMIFEMILSMVLCGGIRNS 223
161 SAKQLIGIVIGIGIAGLTIFGMIFSMVLCALRNS 195

3 9
29DIAX8 PRELIMINARY; PRT; 217 AA.
29DIAX8;
01-JUN-2001 (TREMBLrel. 17, Created)
01-JUN-2001 (TREMBLrel. 17, Last sequence update)
01-OCT-2002 (TREMBLrel. 22, Last annotation update)
3230119D02R1K protein.
3230119D02R1K.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.

```

```

RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann M., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuenh P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirni L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,
RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelincich S., Hill D., Hofmann M., Hume D. A., Kamita M., Lee N. H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombere P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seta T., Shibata Y., Storch K. F.,
RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Haegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK020982; BAB32269.1; -.
DR MGI; MGI:1925095; B230119D02R1K.
DR InterPro; IPR000301; Transmem_4.
DR Pfam; PF00335; Transmembrane4; 1.
DR PRINTS; PR00259; TMF0UR.
DR PROSITE; PS00421; TM4_1; 1.
SQ SEQUENCE 217 AA; 23448 MW; E18DB13DD73055CE CRC64;

Query Match 34.9%; Score 411.5; DB 11; Length 217;
Best Local Similarity 39.7%; Pred. No. 1.9e-31;
Matches 85; Conservative 31; Mismatches 75; Indels 23; Gaps 5;

QY 13 LAGVILGVALMLRHPDPTSLIYELGKRPAPNTFYGYIYLIVAGVMMVFGGCG 72
DB 20 LAGSAVIAFGMLFRGGTMKDL-----SSEDKSPRYFYGLVGVAGALMMVFGGCG 75
QY 73 AIOBQCLLGTFTCLVILFACVAAAGIAGVFNKQDIADVQFTDQALQAVDDDDAN 132
DB 76 AMRBESQCVLGSFFTCILVIFAAVTVGVFAFGKQVAIRHVSMEAVSDYLKDRARG 135
QY 133 AAVVAVKTHETLNCSSNALTTTLTILNSLCPG---GNNLTPLLODCHOKIDELF 189
DB 136 GTLI--THSAFQCCGKSSSQVOPT-----CPKELPG-----HKNCIDKLETVA 179
QY 190 GLVYLIGIAIVAVIMIFEMILSMVLCGGIRNS 223
DB 180 AKQLIGIVIGIGIAGLTIFGMIFSMVLCALRNS 213

RESULT 10
ID 055158 PRELIMINARY; PRT; 235 AA.
AC 055158;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE D6.1A protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BD IX;
RX MEDLINE=98198483; PubMed=9531564;
RA Claas C., Selter S., Claas A., Savelyeva L., Schwab M., Zoeller M.,
RT "Association between the Rat homologue of CO-029, a metabasis-
associated tetraepanin molecular and consumption coagulopathy.";
J. Cell Biol. 141:267-280(1998).
[2]
SEQUENCE FROM N.A.

```


